

SCIENTIFIC REPORTS



OPEN

Alanine substitution in cellobiohydrolase provides new insights into substrate threading

Shigenobu Mitsuzawa^{1,3}, Maiko Fukuura¹, Satoru Shinkawa¹, Keiichi Kimura¹ & Tadaomi Furuta^{1,2}

The glycoside hydrolase family 7 (GH7) member cellobiohydrolase (CBH) is a key enzyme that degrades crystalline cellulose, an important structural component of plant cell walls. As GH7 CBH is a major component in the enzyme mixture used to degrade biomass into fermentable glucose in biorefineries, enhancing its catalytic activity will significantly impact development in this field. GH7 CBH possesses a catalytic tunnel through which cellulose substrates are threaded and hydrolysed. Despite numerous studies dissecting this processive mechanism, the role of amino acid residues in the tunnel remains not fully understood. Herein, we examined the respective contributions of nine amino acid residues in the catalytic tunnel of GH7 CBH from *Talaromyces cellulolyticus* by substitution with alanine. As a result, N62A and K203A mutants were found to possess significantly higher cellulase activities than wild type. Molecular dynamics simulations showed that the N62 residue interacted strongly with the cellulose substrate, impeding threading, while the N62A mutant allowed cellulose to proceed more smoothly. Furthermore, the W63 residue was observed to facilitate twisting of the cellulose substrate in our simulations. This study helps elucidate cellulose threading and provides insight into biomass hydrolysis.

Plant cell wall is a promising feedstock for next-generation biofuels^{1–3}. Its use could diminish the food versus fuel debate associated with conventional biofuels produced from edible biomasses, such as corn starch⁴. Cellulose, the main component in plant cell walls, is a linear molecule composed of glucose units. Cellulases can decompose cellulose into fermentable individual glucose units through hydrolysis. In nature, mixtures that comprise several tens of different enzymes secreted by some fungi and bacteria synergistically are used to hydrolyse recalcitrant crystalline cellulose. In contrast, conventional starch-based biofuel production uses only a few enzymes for biomass hydrolysis⁵. The recalcitrance of cellulose is a critical obstacle to the commercialisation of cellulosic biofuel. In a cost estimation of ethanol production from corn stover, cellulase in substrate hydrolysis contributed \$0.34/gal of ethanol product to the total cost of \$2.15/gal⁶, which is almost eight times higher than the cost of amylase used in ethanol production from shelled corn⁷. Therefore, enhanced cellulase activity, which would lead to reduced enzyme costs, is greatly needed.

As enzyme cost is the product of the amount used and the production cost per unit amount, the enzyme cost can be reduced by separately lowering these two factors. Previously, we constructed synthetic mixtures consisting of three cellulase components from *Talaromyces cellulolyticus* and two additive hemicellulases from different organisms^{8,9}. *Talaromyces cellulolyticus* is a cellulose-degrading fungus originally isolated in Japan^{10–12} and regarded as a promising major source of commercial cellulases¹³. This defined enzymatic cocktail gave a higher sugar yield than that of the culture supernatant of *T. cellulolyticus* at the same dosage. Therefore, a targeted sugar yield can be obtained with a lower amount of enzyme using the synthetic cocktail than the native mixture. To reduce production costs, we constructed *Aspergillus oryzae* strains that were genetically modified to overproduce these enzyme components¹⁴. To further decrease the amount of enzyme used for saccharification, we aimed to improve the components in the cocktail. In the synthetic cocktail, the glycoside hydrolase family 7 (GH7) member cellobiohydrolase (CBH) made up 47% of the total enzyme content in the optimised composition¹⁵. GH7 CBH is also the most abundant component in typical cellulolytic enzyme systems expressed by fungi¹⁶. Therefore,

¹Fundamental Technology Research Center, Honda R&D Co., Ltd., 1-4-1 Chuo, Wako-shi, Saitama, 351-0113, Japan. ²School of Life Science and Technology, Tokyo Institute of Technology, B-62 4259 Nagatsuta-cho, Midori-ku, Yokohama, 226-8501, Japan. ³Present address: Honda Research Institute Japan Co., Ltd., 8-1 Honcho, Wako-shi, Saitama, 351-0188, Japan. Correspondence and requests for materials should be addressed to S.M. (email: shigenobu.mitsuzawa@jp.honda-ri.com)

we resolved to improve this enzyme using molecular engineering before focusing on other components. The amount of enzyme required for saccharification is affected by certain properties of the enzyme species employed, including specific activity, degree of product inhibition, and thermal stability. Of these properties, specific activity is the most dominant. Therefore, in this study, we have attempted to enhance the specific activity of GH7 CBH from *T. cellulolyticus* (*TcCel7A*).

GH7 CBH from *Phanerochaete chrysosporium* (*PcCel7C*) has previously been genetically modified¹⁷. *PcCel7C* variants with single site-directed mutations were constructed on and near the catalytic tunnel using an *in vitro* expression system, and demonstrated that some mutants possessed higher activities than those of wild type (WT) toward phosphoric acid swollen cellulose (PASC) and 4-methylumbelliferyl β -D-cellobioside (MUC)^{18,19}. However, the molecular mechanism for the effect of amino acid mutations on enzymatic activity has yet to be elucidated. The molecular structure and catalytic mechanism of GH7 CBH from *Trichoderma reesei* (*TrCel7A*) have been studied most intensively^{20,21}. GH7 CBH cleaves cellobiose at the reducing end of cellulose substrates sequentially as it slides along the cellulose surface through a processive mechanism^{22,23}. Upon enzyme attachment to the surface of crystalline cellulose, the polysaccharide chain initiates threading into the catalytic tunnel. The enzyme then successively repeats the cycle of hydrolysis, product expulsion, and threading²⁴. Therefore, threading is a key process in the processive mechanism determining the efficiency and reaction rate of overall degradation^{25,26}. The driving force behind this process has been a topic of intense investigation. Based on the crystal structure of *TrCel7A*, the catalytic tunnel is 50 Å long and contains nine glucosyl binding subsites (−7 to +2), with catalytic sites located around subsites −1 and +1²⁷. The interaction between the substrate and the long catalytic tunnel is considered to govern both stepwise movement and hydrolytic activity^{28,29}. The strongest interaction has been suggested to occur at the expulsion site (subsites +1 and +2)^{30,31}, while the next most important location is the first half of the tunnel (subsites −7 to −3)²⁹. Subsites −7 to −3 are aligned by hydrogen bond-forming amino acid residues and tryptophan residues, providing hydrophobic stacking interactions. Those residues are thought to guide the cellulose strand toward the catalytic site. To achieve a smooth processive movement, these interactions should be fine-tuned. The importance of the tryptophan residue (W40) at the entrance of the tunnel (subsite −7) of *TrCel7A* in guiding cellulose strands into the catalytic tunnel has been confirmed by replacing the residue with alanine³². Furthermore, the tryptophan residue (W38) near subsite −4 has been replaced with alanine to elucidate the rate-limiting step of *TrCel7A*³³. Aromatic residues that line the catalytic tunnels of chitinase, another processive glycoside hydrolase, have also been shown to affect processivity by mutating their residues to alanines^{34,35}. However, a comprehensive understanding of interactions between glucosyl units situated at subsites and amino acid residues along the tunnel has yet to be obtained. In contrast to tryptophan residues, hydrogen-bond-forming residues have received relatively little attention regarding their effect on threading. Although molecular simulations on *TrCel7A* have demonstrated that electrostatic interactions with conserved polar residues at subsites +2 and −1 provide the thermodynamic driving force behind cellulose chain translocation³⁶, other polar residues on the catalytic tunnel have not been investigated. Kostylev *et al.*³⁷ reported no significant change in catalytic activity when asparagine and threonine residues at the tunnel entrance of a processive GH48 exocellulase from *Thermobifida fusca* (*TfCel48A*) were mutated to alanines.

In this study, we conducted alanine scanning experiments on *TcCel7A*³⁸ for nine amino acids located along subsites −7 to −3 with the aims of obtaining mutants with higher specific activities and understanding the role of these residues in the enzyme catalytic mechanism. Wild type and mutant *TcCel7A* were subjected to activity assays with commercially available microcrystalline cellulose and pretreated corn stover. Furthermore, we performed molecular dynamics simulations for WT and a selected alanine-mutant of *TcCel7A* to further elucidate the molecular mechanism.

Results

Structure of *TcCel7A*. First, we constructed a homology model for *TcCel7A* using *TrCel7A* (PDB:8CEL²⁷) as the template (Fig. 1a, Supplementary Movie 1) because the structure of this enzyme has yet to be solved. The amino acid sequence of *TcCel7A* showed high similarity with that of *TrCel7A* (64% identity and 90% similarity) (Supplementary Fig. S1) and the resultant model structure of *TcCel7A* was very similar to the structure of *TrCel7A*, with a root mean square deviation (RMSD) between main chain atoms of 1.06 Å (Supplementary Fig. S2). We next conducted a preliminary 25-ns MD simulation of *TcCel7A* in water. The resultant RMSD from the initial structure was almost constant, at around 2 Å, throughout the time evolution, indicating that the conformation of the enzyme–substrate complex was stable (Fig. 1b). Furthermore, the distances between each subunit and the catalytic residue (E234) were almost constant (Fig. 1c).

TcCel7A possessed a tunnel-like structure with four loop regions surrounding the substrate cellulose chain, which is typical of GH7 CBH²⁶, although the upper left loop, shown in Fig. 1a, was three amino acids shorter than that in *TrCel7A* (Supplementary Figs S1 and S2, Loop 1). In the model structure (Fig. 1a), the cellulose chain had an extended conformation at subsites −7 to −4 that was relatively close to that of crystalline cellulose. However, the cellulose chain was then radically twisted at subsites −3 and −2 prior to the catalytic site (−1/+1). This feature suggested that amino acid residues along the tunnel interacted with and applied torsional force to the substrate chain, propelling it further into the catalytic tunnel. This characteristic conformation of the cellulose chain was also confirmed during the above MD simulation (Supplementary Fig. S3).

Among these structural features, we concentrated on subsites −7 to −3 in this study, examining interactions between amino acid residues and the substrate. Therefore, we selected nine amino acid residues in this region for alanine scanning experiments. These amino acids were close (within 4 Å) to subsites −7 to −3 and were classified into three groups according to their location in the molecular structure: Entrance (Fig. 2a), Base (Fig. 2b), and Loop (Fig. 2c). Interactions of the nine alanine residues with subsites are summarised in Fig. 2d.

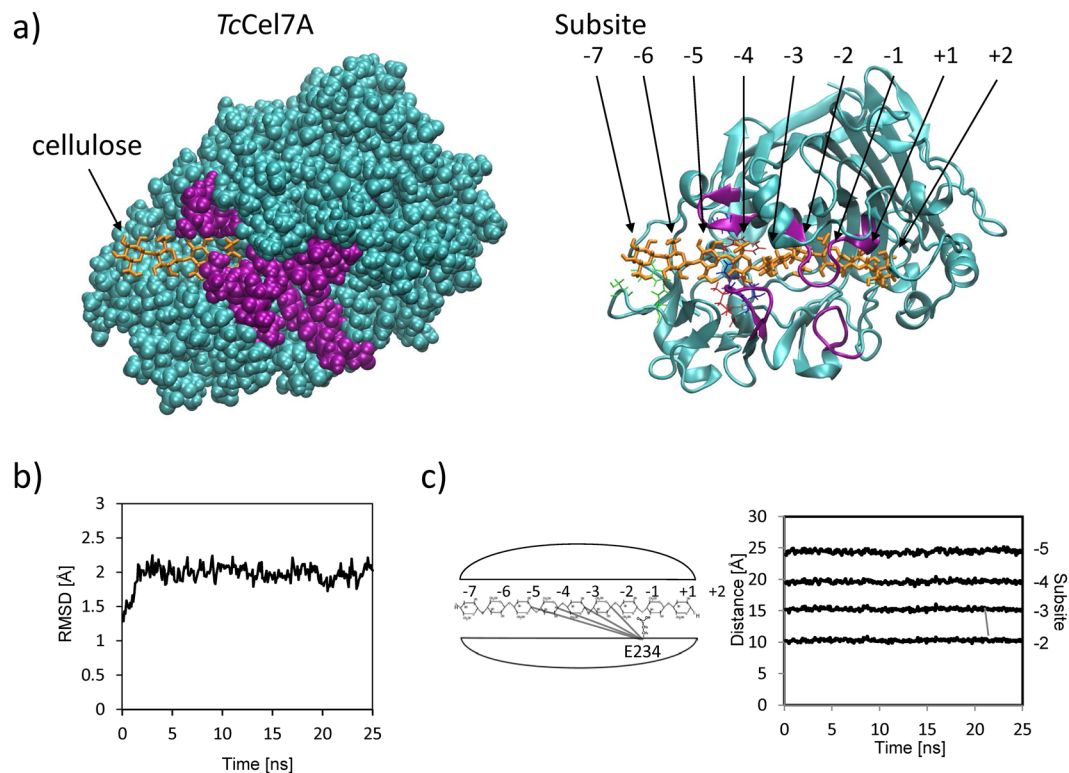


Figure 1. Model structure of *TcCel7A* containing cellulose in the catalytic tunnel and its MD simulation. **(a)** Overviews in VDW presentation (left) and in ribbon presentation (right). Four loop regions surrounding the substrate cellulose chain are depicted in purple. **(b)** Time evolution of RMSD from the initial structure and **(c)** of distances between each subunit and the catalytic residue (E234).

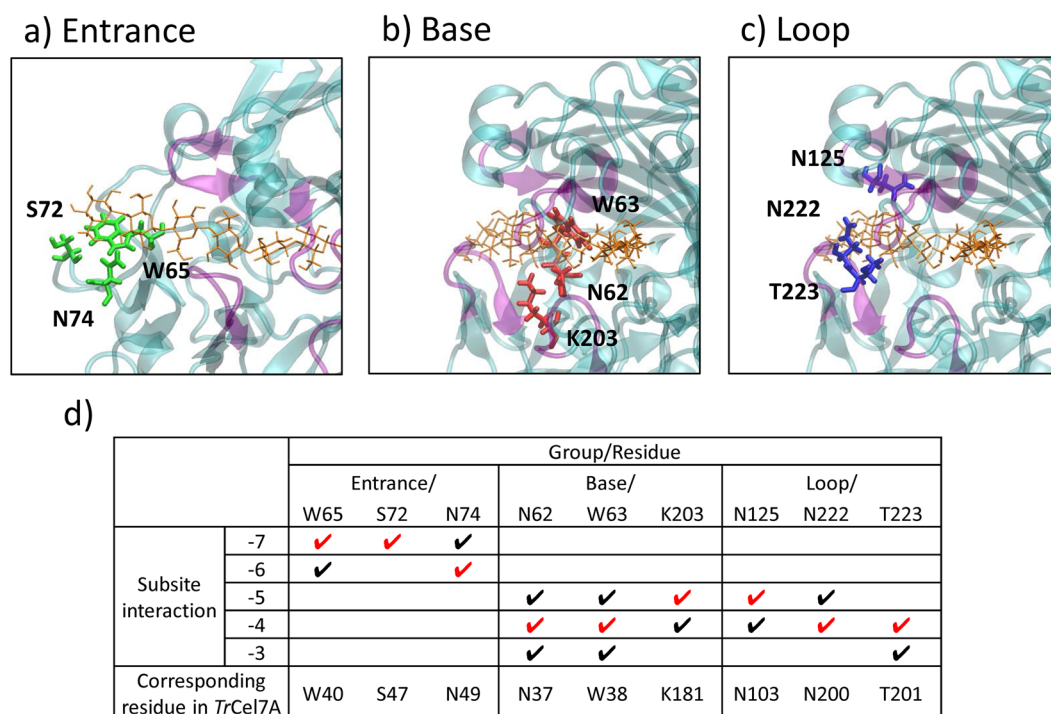


Figure 2. Nine amino acid residues used in the alanine scanning experiment at three locations: **(a)** Entrance, **(b)** Base, and **(c)** Loop. **(d)** Interactions of the nine alanine-scanning residues with subsites. Interactions are indicated by check marks, with the main interactions shown in red for each residue. The corresponding residues in *TrCel7A* are listed.

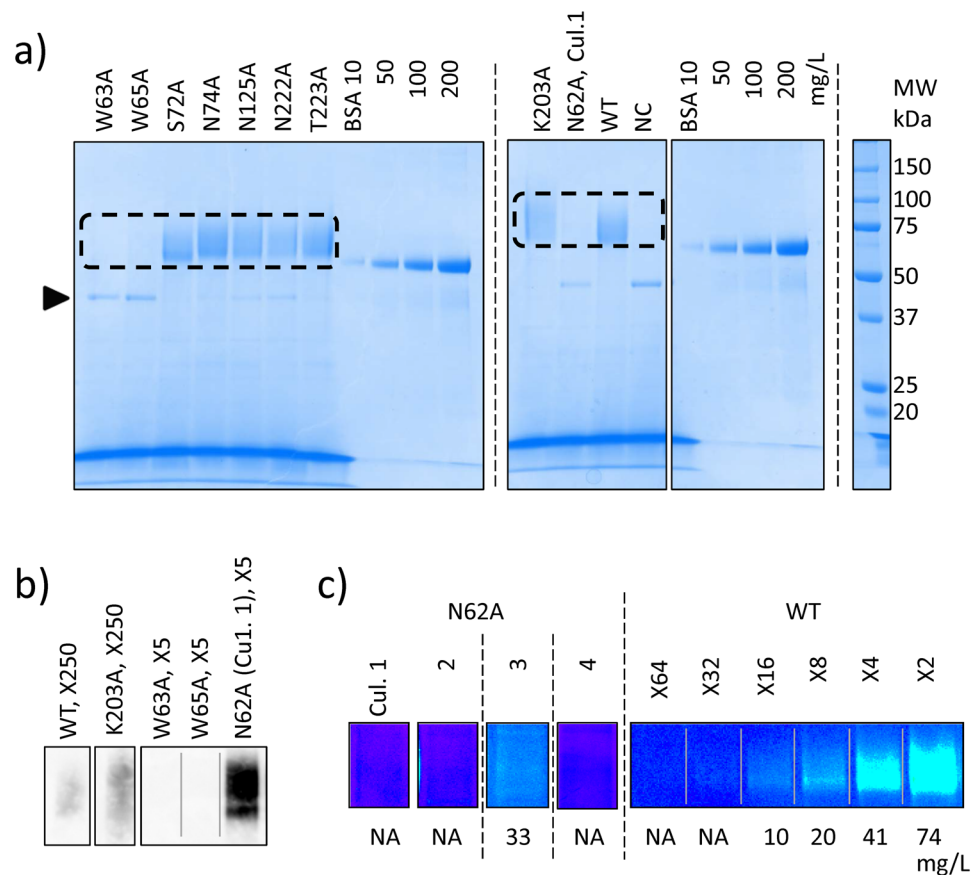


Figure 3. Production of wild-type (WT) and nine alanine-substituted mutants of *TcCel7A*. (a) SDS-PAGE of WT and nine mutants. NC: untransformed *A. oryzae* strain NS4DLP. BSA concentrations: 10, 50, 100, and 200 mg/L. Target CBHs are shown in a dashed square and the α -amylase position is denoted with a wedge. Full-length gels are presented in Supplementary Fig. S2. (b) Western blotting of WT, K203A, W63A, W65A, and N62A. Dilution ratios follow the name for each sample. All images are excerpts from an identical membrane (Supplementary Fig. S3). (c) Pseudo-coloured SDS-PAGE images of N62A from four different cultures and of WT with serial dilutions ($\times 2$ –64) for comparison. For each lane, concentration was determined using image densitometry. Full-length gels with original colour are presented in Supplementary Fig. S4.

Production of *TcCel7A* proteins. Prior to this study, we first attempted to express *TcCel7A* WT in four microorganism species, namely *Escherichia coli*, *Bacillus brevis*, *Pichia pastoris*, and *A. oryzae*, with only *A. oryzae* able to produce the protein. Therefore, in this study, WT and nine alanine-substituted mutants of *TcCel7A* were produced by employing *A. oryzae* strain NS4DLP as an expression host (Fig. 3). The production of each target protein was assayed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of the culture supernatant (Fig. 3a, Supplementary Fig. S4). WT and six mutants (S72A, N74A, N125A, N222A, T223A, and K203A) were clearly recognisable in relative abundance. In contrast, W63A, W65A, and N62A mutants were not detected as visible bands. When treated with PNGase F, *TcCel7A* WT showed a narrower band on SDS-PAGE (Supplementary Fig. S5), suggesting that the recombinantly produced protein was *N*-glycosylated by the *A. oryzae* host³⁹. It was confirmed that *N*-glycosylation did not affect the weight specific activity in comparison of the WT's, purified from the culture supernatant of *T. cellulolyticus* and recombinantly produced by *A. oryzae*. To examine the production of W63A, W65A, and N62A mutants, we conducted western blotting (Fig. 3b, Supplementary Fig. S6). WT and K203A were also assayed as positive controls. Based on the gel, we confirmed N62A production, but not that of W63A or W65A.

In a preliminary trial to treat mutant proteins including K203A and N62A with ultrafiltration (centrifugal concentration/purification), loss of activity was found that was not observed with WT. In this study, with the aim of obtaining new insight into substrate threading rather than rigorously characterising the highly active mutants, we characterised the mutant proteins only as the culture supernatant of production host strain without any concentration or purification treatments. As the culture supernatant samples contained several proteins other than *TcCel7A*s and the fragmental amino acids/peptides, as shown Fig. 3a, quantification of the target proteins was carried out using SDS-PAGE image data. In comparison with the serial dilution of WT on SDS-PAGE gel, we determined the concentrations of four cultures of N62A (Fig. 3c, Supplementary Fig. S7). The yields of cultures 1, 2, and 4 were below 5 mg/L. For culture 3, N62A was discernible on the SDS-PAGE gel, allowing the concentration to be determined (33 mg/L, Supplementary Fig. S8). Table 1 summarises the secreted amounts of N62A (Culture 1) and other target proteins.

	NC	WT	Entrance/			Base/			Loop/		
			W65A	S72A	N74A	N62A (Culture 1)	W63A	K203A	N125A	N222A	T223A
Production [mg/L]	<5	149	<5	171	229	<5	<5	117	168	126	213
Cellobiose	ND	0.778 (0.01)	ND	0.759 (0.016)	0.860 (0.026)	0.100 (0.013)	ND	0.768 (0.033)	0.882 (0.037)	0.603 (0.032)	0.703 (0.044)
Activity	ND	1.00 (0.01)	ND	0.851 (0.02)	0.720 (0.02)	>3.85 (0.49)	ND	1.26 (0.05)	1.01 (0.04)	0.908 (0.05)	0.634 (0.04)

Table 1. Production and hydrolytic activity of WT and nine alanine-substituted mutants of *TcCel7A*. Production refers to the concentration (mg/L) of each enzyme obtained from the SDS-PAGE image in Fig. 3a. Cellobiose refers to the HPLC peak area of cellobiose (arbitrary unit) delivered by *TcCel7A* in the activity assay. Activity refers to hydrolytic activity, derived from dividing the Cellobiose value by the Production value. Average Cellobiose and Activity values are listed with standard deviations in parentheses. ND: Not detected. Higher activities are shown in bold (N62A and K203A).

Hydrolytic activity. From secretome analysis by SDS-PAGE (Fig. 3a), we found that, other than CBH, the culture supernatant contained a single detectable band identified as α -amylase of the *A. oryzae* host from its molecular weight⁴⁰. Using the culture supernatant of the untransformed strain (NC), we confirmed that this protein exhibited no cellulase activity on Avicel, carboxymethylcellulose (CMC), or 4-nitrophenyl β -D-glucopyranoside (pNPG). Therefore, we conducted the following experiments employing these culture supernatants (0.22- μ m filtrated, but unpurified) as CBH samples.

First, we performed hydrolysis experiments for WT and the nine mutants using commercially available microcrystalline cellulose as the substrate. Cellobiose was detected as the product by HPLC (Fig. 4a, Supplementary Fig. S9). The specific weight activity was calculated for each enzyme species relative to WT by dividing the cellobiose amount by the enzyme amount (Table 1). The benchmark value of WT was 0.204 U/mg. Among the nine mutants, N62A and K203A exhibited significantly higher activities than WT. As the N62A concentration was at least 30 times lower than that of WT, we also examined 1/10 and 1/100 dilutions of WT (Fig. 4b). A comparison of N62A (measured) and WT (estimated) at the same concentration (<0.56 mg/L, stock concentration of below 5 mg/L divided by the dilution when preparing the reaction mixture) confirmed that N62A was at least twice as active as WT.

The hydrolytic activities of WT and the two mutants were also compared using different assay methods and on different substrates, including cellulose from Merck and pretreated corn stover. Data showed that N62A had more than two-fold greater activity than WT on all substrates (Table 2). The enhancement exhibited by N62A might include some overestimation due to the lower test concentration compared with that of WT, as discussed above.

Molecular dynamics (MD) simulation. To elucidate the mechanism behind the activity enhancement afforded by replacing N62 with alanine, we conducted MD simulations of WT and the N62A mutant (two runs each, Fig. 5, Supplementary Movies 2 and 3). Starting from the initial position of the reducing end (cellononoase) at subsite -5 (Fig. 5a), we examined the processive movement of the cellulose strand into the catalytic tunnel. For motion analysis, we measured the distance between the C1 atom of the reducing end glucose and C₆ atom of E234 at the catalytic site (Fig. 5b,c). Snapshots at four different stages along the time course are shown in Fig. 5d. For both runs using WT, the cellulose strand had almost proceeded to subsite -3 after around 15 ns. Notably, when the reducing end glucose had proceeded to subsite -4 (1st, 11.5 ns), it was twisted and aligned along W63, where the two ring planes were almost parallel with a distance of about 4 Å, suggesting a weak hydrophobic stacking interaction²⁷. After the cellulose strand reached subsite -3 (1st, 15.0 ns), it was observed to move backwards, finally fluctuating between subsites -4 and -5 (1st, 124.8 ns). In this state, the hydroxyl group at C6 in the leading glucosyl unit and the sidechain of N62 were in close proximity (distance between hydroxyl group oxygen atom and N62 oxygen atom was 2.6 Å; for the time course, see Supplementary Fig. 10a), suggesting a tight hydrogen bond. On the other hand, in one run using N62A (Fig. 5e), we observed twisting of the reducing end glucose in association with W63 (1st, 32.0 ns), but no apparent recession of the cellulose strand. The cellulose strand proceeded smoothly toward subsite -3 and stayed there (1st, 109.0 ns). In the other run, the distance stayed at around -4 during the time course. However, close examination of the typical conformation of the reducing end glucose (2nd, 99.5 ns) showed that it was almost identical to that of the cellulose staying at subsite -3 in the previous run (1st, 109.0 ns).

Next, we compared the interactions of N62 and subsite -4 with that of N222 and subsite -4 in WT simulations. Notably, the two asparagine residues interacted with subsite -4 (Fig. 6a). For the respective asparagine residue, histograms of the dihedral angle formed by N-C_α-C_β-C_γ in “threading” (Fig. 5b, 60–160 ns trajectory in the first run (black line); for the time course, see Supplementary Fig. 10b) or “occupied” (Fig. 1b) cases are shown in Fig. 6b,c. N62 adopted almost the same conformation, with a dihedral angle of approx. 300° (main frequency, >45%), in the simulation for the occupied case. In contrast, in the threading case, a similar conformation had a frequency of 25% (major, but less than in the occupied case, see Fig. 6d for a typical conformation), while the twisted conformation with dihedral angles <220° made a significant distribution (minor) that was not observed in the occupied case. This conformation was favourable for hydrogen bonding with the hydroxyl group at C6 of the leading glucosyl unit positioned between subsites -4 and -5, as previously shown in Fig. 5d, at 124.8 ns (dihedral angle = 174°). In the N222 reference, the positions and shapes of the three peaks were almost identical in the occupied and threading runs, although some changes in the frequency distribution among the three peaks were observed.

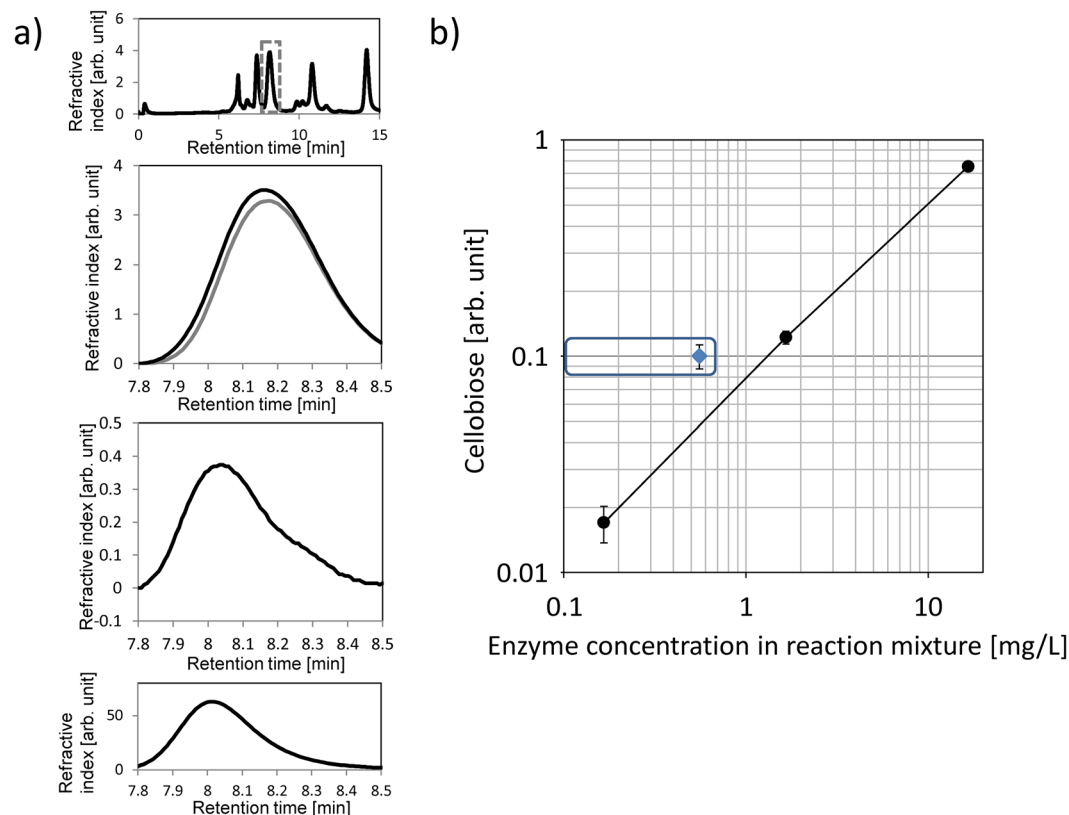


Figure 4. Hydrolytic activity assays for N62A mutant. **(a)** Chromatograms of reaction aliquots with N62A and cellobiose standard. Upper two charts show the retention time ranges 0–15 min and 7.8–8.5 min. Black line: reaction aliquot with substrate. Grey line: reaction aliquot without substrate. Third chart shows the difference in the two values (with and without substrate). Fourth chart shows the chromatogram of cellobiose, 0.15% (w/v). **(b)** Hydrolytic activity of WT at three different concentrations and the N62A mutant. Cellobiose refers to the HPLC peak area of cellobiose delivered by the action of *TcCel7A* in the activity assay. Black circles: WT. Blue diamond: N62A mutant, assuming a concentration below 0.5 mg/L in the reaction mixture. Error bar: Standard deviation.

Substrate	Assay	Activity			Culture no. of N62A
		WT	N62A	K203A	
Cellulose, Sigma (excerpted from Table 1)	HPLC	1.00 (0.01)	>3.85 (0.49)	1.26 (0.05)	1
Cellulose, Merck	HPLC	1.00 (0.04)	>4.57 (0.45)	1.36 (0.05)	
Cellulose, Merck	HPLC	1.00 (0.05)	>3.78 (0.90)	ND	2
Cellulose, Sigma	DNS	1.00 (0.42)	2.20 (1.12)	1.89 (0.58)	3
Pretreated corn stover	HPLC	1.00 (0.07)	>3.92 (0.34)	1.00 (0.15)	4

Table 2. Hydrolytic activities of WT, N62A, and K203A from different culture batches on different substrates. Hydrolytic activity values from duplicated experiments are listed with standard deviations in parentheses.

Moreover, to explore the free energy landscapes of substrate threading in WT and the N62A mutant, we calculated the binding free energies of the cellulose substrate with WT or N62A at each subsite using the molecular mechanics/generalized Born surface area (MM/GBSA) method using the trajectories of the occupied simulations (Fig. 6e). The landscapes shared a common funnel-like topology that had a global minimum point at subsite –1 (catalytic site), a local minimum point at subsite –4 (N62 hydrogen bonding), and a local maximum point at subsite –3. At the local maximum point (subsite –3), the binding free energy for N62A was slightly weaker than that for WT (by 0.75 kcal/mol), which was attributed to the lack of hydrogen bonds with the N62 residue.

Discussion

In this study, to investigate the structural features of the catalytic tunnel of GH7 CBHs, homology modelling of *TcCel7A* was conducted using *TrCel7A* as a template (Fig. 1a, Supplementary Fig. S2, Supplementary Movie 1). The amino acid sequence and resultant structure of *TcCel7A* were compared with those of *TrCel7A* and *PcCel7D* (Supplementary Figs S1 and S2), showing that the three proteins shared similar primary and tertiary structures. In

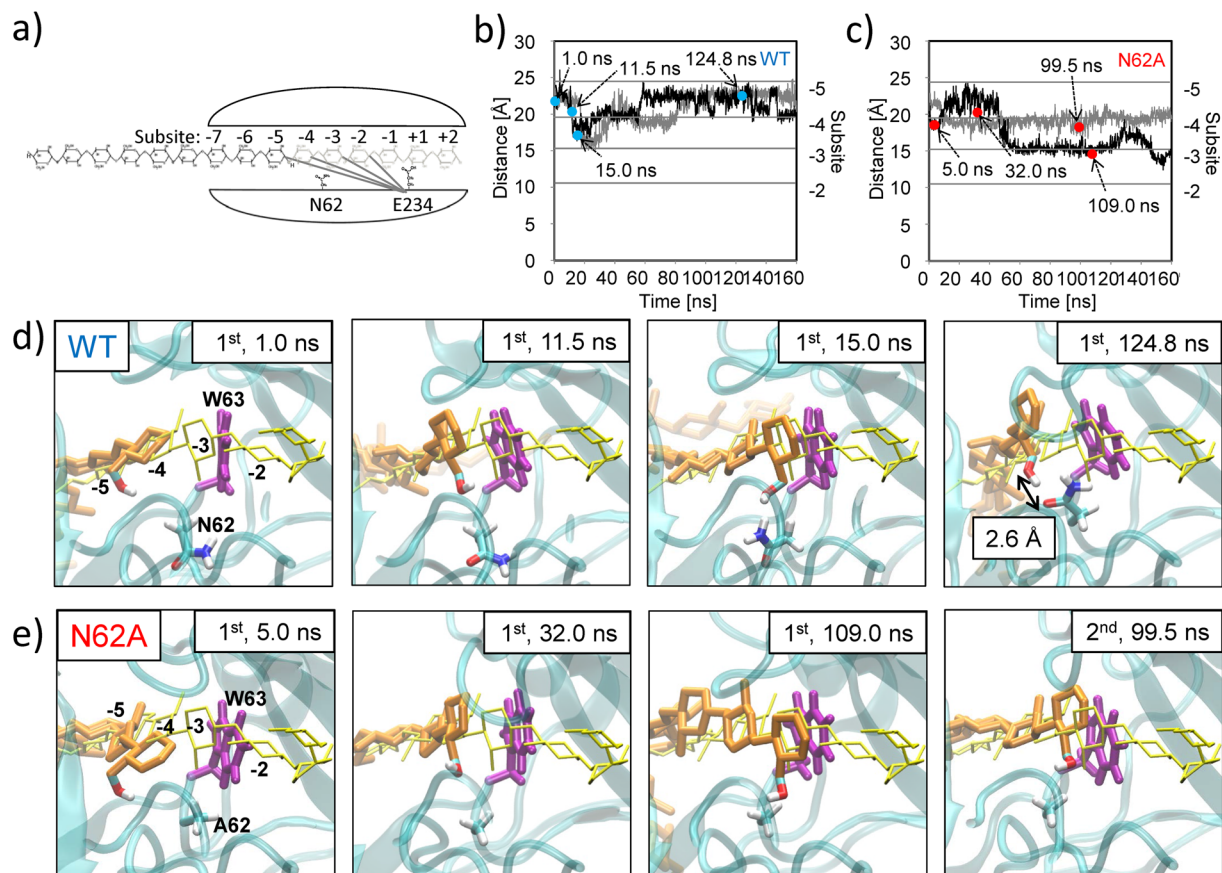


Figure 5. Substrate threading in MD simulations of WT and the N62A mutant. (a) Schematic diagram of the initial cellulose position (reducing end at subsite -5) and the distance between each subsite (-5 to -2) and catalytic residue E234. (b) Time evolution of distance between the C1 atom of the reducing end glucose and C $_{\alpha}$ atom of catalytic base E234 in the MD simulation of WT (first run in black, second run in grey), and (c) the same time evolution in the MD simulation of N62A (similar colours). The distances for each subsite (-5 to -2) in (a) are plotted with a line to scale the position of the cellulose substrate. Several stages are indicated by blue (WT) or red (N62A) circles. (d) Snapshots of substrate threading in the MD simulations of WT depicted at 1.0, 11.5, 15.0, and 124.8 ns from the first run correspond to initial (1.0 ns), pre-procession (11.5 ns), procession (15.0 ns), and recession (124.8 ns) stages, respectively. The threading cellulose chain is represented by its constituent carbon/oxygen ring chain (thick orange stick). In the leading glucosyl unit, the hydroxyl group at C6 is also drawn. To indicate the position of each subsite, the cellulose strand from the homology model (Fig. 1a) is superimposed (thin yellow stick of carbon ring). N62 and W63 sidechains are shown as thick sticks (CPK colouring and purple, respectively). (e) Substrate in the N62A mutant depicted at 5.0, 32.0, and 109.0 ns from the first run and 99.5 ns from the second run, corresponding to initial (1st, 5.0 ns), pre-procession (1st, 32.0 ns), and procession (1st, 109.0; 2nd, 99.5 ns) stages, respectively.

addition to the structure of the protein, the cellulose substrate had a characteristic conformation when complexed with GH7 CBHs. In the subsequent MD simulation, the model structure of *TrCel7A* was confirmed to be very stable, as evaluated using the RMSD (Fig. 1b) and the distances between each subunit of the cellulose substrate and the catalytic residue (E234) (Fig. 1c). Moreover, we confirmed the constancy of cellulose chain conformation, especially at subsites -3 to -5 (Supplementary Fig. S3).

By employing *A. oryzae* as an expression host, we succeeded in producing mutants, except for tryptophan-to-alanine mutants, namely W63A and W65A (Fig. 3). Previously, Nakamura *et al.*³² obtained the corresponding W65A mutant of *TrCel7A* (W40A) by employing *T. reesei* as the expression host. Moreover, Kari *et al.*³³ obtained the corresponding W63A mutant of *TrCel7A* (W38A) by employing *A. oryzae* as the expression host. Although the reason for failed expression of these two mutants of *TcCel7A* in the present system was not clear, a possible explanation was that W63 and W65 are indispensable in maintaining the molecular structure of *TcCel7A*. Whether the failure to produce these alanine-substituted mutants depended on the synthetic genes and/or expression host would be resolved via expression studies of *TrCel7A* mutants in *A. oryzae* and/or of *TcCel7A* mutants in *T. reesei*. We found that *TcCel7A*s recombinantly expressed by *A. oryzae* was highly *N*-glycosylated. This *N*-glycosylation did not affect the weight specific activity for WT. However, its effects on the productivity or stability of recombinantly expressed *TcCel7A* are yet to be studied. As some mutant proteins were unstable in ultrafiltration (centrifugal concentration), we decided to characterise the mutant proteins only as the culture

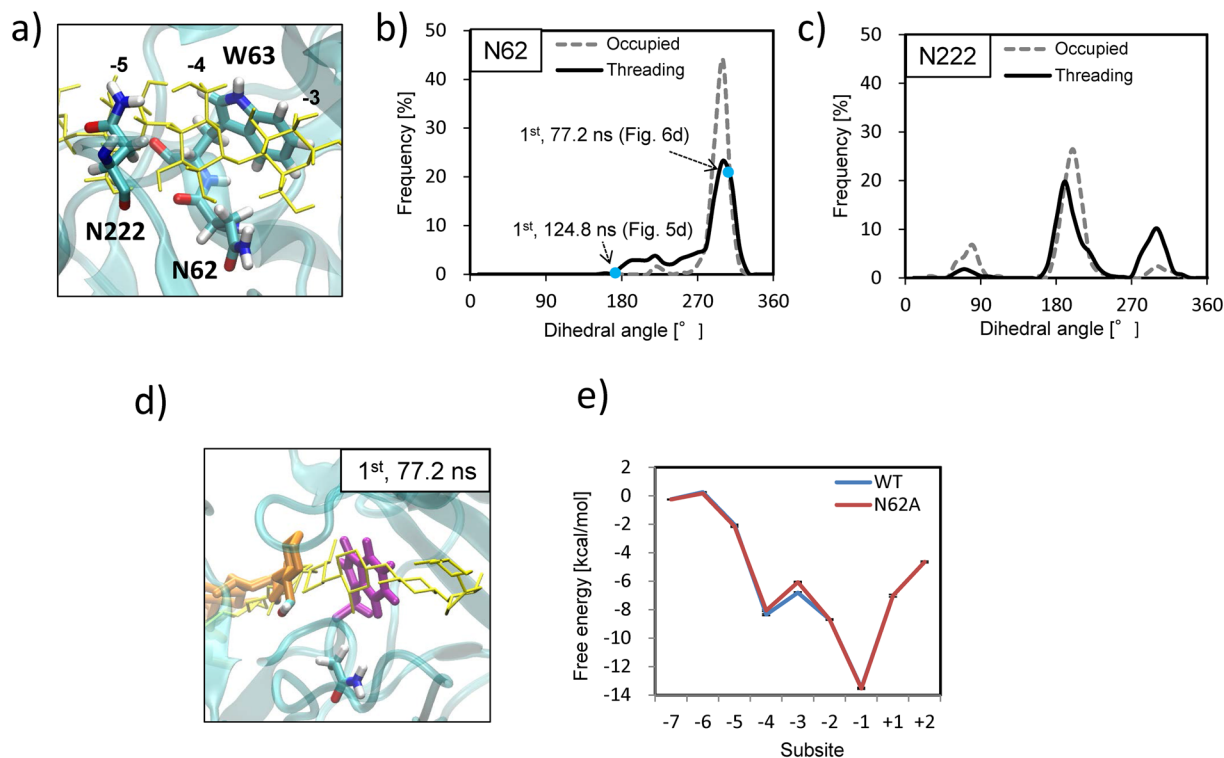


Figure 6. Comparison of the interactions of N62 and subsite -4 with that of N222 and subsite -4 in WT simulations. **(a)** Subsites -3 to -5 and three residues near the location: N62, W63, and N222. **(b)** Histograms of the dihedral angle ($N-C_{\alpha}-C_{\beta}-C_{\gamma}$) of N62 in substrate threading (black solid) and occupied (dashed grey) simulations, and **(c)** the same histograms for N222 (similar colours). **(d)** Typical conformation of the major frequency with a dihedral angle of 309° (approx. 300°) for N62 (1st, 77.2 ns). **(e)** Free energy landscapes of substrate threading in WT and the N62A mutant. Binding free energy of the cellulose substrate with WT (blue) or N62A (red) mutant at each subsite was calculated using the MM/GBSA method. Error bar: Standard error.

supernatant of the production host strain without concentration or purification. Other purification/concentration methods, such as employing affinity tags, require testing to further examine the reaction kinetics with *TcCel7A* WT and mutants.

High CBH loading on cellulose substrate is known to lead to substrate saturation, where the activity is not proportional to enzyme loading⁴¹. In this study, the substrate concentration was 4% (w/v) and the enzyme loading was around 10 mg/L or less (Fig. 4b), which corresponded to less than $1 \mu\text{M}$. This enzyme loading was in the range where activity is proportional based on the *TrCel7A* data⁴¹. The activity data obtained with *TcCel7A* WT and its 1/10 and 1/100 dilutions confirmed this assumption (Fig. 4b). More extensive assays will be considered in future investigations, preferably using purified WT and mutants, to study the dependence of activity on enzyme loading and more precisely determine the weight specific activity enhancement ascribed to the amino acid mutation.

The Avicel assay of the alanine mutants (Table 1) showed that mutations at the Entrance or Loop moderately decreased the activity, except for N125A, in which the activity was unaffected, and W65A (not detected), while those at the Base increased the activity, except for W63A (not detected). The decreased activity in the former case could be explained by the conventional knowledge that a processive enzyme might bind its polymeric substrate with moderate affinity at multiple adjacent sites distributed within a catalytic tunnel²². Therefore, the contribution of the processive mechanism from each interacting residue was relatively weak, with one alanine mutation not leading to a crucial loss in cellulase activity. In contrast, the increased activity in the latter case was somewhat unexpected. Among the two mutants possessing significantly higher activity than that of WT toward microcrystalline cellulose, the N62A mutant showed higher activity than WT toward pretreated corn stover, while the enhanced activity of the K203A mutant was diminished toward this natural biomass. It could be postulated that the enhanced activity of the K203A mutant toward reagent cellulose was due to a lower substrate affinity, which is supposedly prominent when the specific cellulose surface area accessible to the enzyme is larger³³.

To elucidate the mechanisms underlying the activity enhancement in the N62A mutant, we performed MD simulations of initial threading, which demonstrated distinctive time-evolving trajectories for WT and N62A (Fig. 5, Supplementary Movies 2 and 3). The most prominent finding was that the leading glucosyl unit deviated from the correct path and was sterically hindered from moving further by N62 (Fig. 5d, 1st, 124.8 ns). The histogram of the dihedral angle of N62 during threading (Fig. 6b) showed that this conformation was unique to “threading” and not seen in the “occupied” cases. N62 appeared to play a role in pulling the cellulose substrate into the catalytic tunnel by adopting the conformation with a smaller dihedral angle, which could have a positive

effect on processive movement. However, this temporary hydrogen bonding also tended to deviate the cellulose substrate from the correct path, as described above. Our activity assay demonstrated that this negative effect outweighed the positive effect, which was related to recession of the threading. Substitution of N62 with alanine resulted in a smoother surface inside the catalytic tunnel with no protruding and interacting residues, which led to a higher activity. In the reference case, N222 did not exhibit any characteristic conformation associated with substrate threading. Two MD simulation runs were performed for N222A, in which no processive movements of the substrate cellulose were observed (Supplementary Fig. S11).

When the leading glucosyl unit had proceeded to subsite -4 in the MD simulations, W63 was found to facilitate twisting of the leading glucose, possibly due to a weak hydrophobic stacking interaction, so that the leading glucose could easily proceed to subsite -3 (Fig. 5d, 1st, 11.5 ns). Although the crystal structure suggests that the W63 homologue interacts with the glucose at subsite -4 of the cellononase complexed with *TrCel7A*²⁷, the major twisting of the cellulose chain at subsite $-4/-3$ was not ascribed to this interaction. To our knowledge, our results are the first demonstration of how W63 functions dynamically as the substrate proceeds through the catalytic tunnel. When the leading glucosyl unit was twisted at subsite -4 , its C6 moiety was situated on the opposite side to cellononase in the homology model and was, therefore, in close proximity to N62, resulting in hydrogen bonding (Fig. 5d, 1st, 124.8 ns). N62 was likely to have a similar effect on the successive threading of cellobiose per cycle. In a trial MD simulation, starting from an initial configuration in which subsites -5 to -1 were occupied with a cellulose strand and subsites $+1$ and $+2$ were empty, no significant substrate movement along the catalytic tunnel was observed within the calculation time of 160 ns.

The W38A mutant of *TrCel7A* (corresponding to the W63A mutant of *TcCel7A*) has been shown to produce a two-fold increase in the maximum quasi-steady-state hydrolytic rate in a certain range of substrate and enzyme loading, suggested to be caused by reduced substrate affinity³³. More comprehensive data are required to better understand the function of W38.

The free energy landscape of substrate threading is shown in Fig. 6e, in which the energy contributions are resolved for each glucosyl unit. The landscape shape was funnel-like, with the global minimum at subsite -1 . This seemed reasonable for the enzyme to repeatedly perform the cycle of substrate chain threading, hydrolysis, and cellobiose expulsion. Hydrolysis took place at subsite $-1/+1$, at which point the substrate was rigidly fixed to the correct coordinate by interactions with the surrounding residues. This was in good accordance with previous findings that the fluctuation of the glucosyl unit at subsite -1 was smaller than those at the entrance (subsites -7 and -6) or end (subsite $+2$) of the catalytic tunnel, as demonstrated in the MD simulations of GH7 CBHs from different organisms^{28,42}. Notably, at subsite -3 , the binding energy reached a local maximum, suggesting a loose interaction between the substrate and enzyme in that vicinity. At subsite -3 , the binding free energy of N62A was weaker than that of WT by 0.75 kcal/mol, roughly equivalent to the energy of one hydrogen bond (0.5–1.8 kcal/mol)⁴³. Colussi *et al.*²⁹ conducted calorimetric experiments to study the interactions of cello-oligosaccharides with a catalytically deficient mutant (E212Q) of *TrCel7A*, and suggested that filling subsites -2 and -1 had no, or a slightly unfavourable, impact on the net affinity (ΔG°), despite these subsites being rich in enzyme-ligand hydrogen bonding. Therefore, these stabilising bonds were compensated by unfavourable steric effects, including cellulose strand twisting or pyranose ring distortion. Substrate twisting was not explained by the static image of interaction energy and, therefore, might be more accurately interpreted as a dynamic phenomenon, including the contribution of the temporal action of the surrounding residues, such as N62 and W63 herein. The loose and appropriate interaction between the glucosyl unit and enzyme residues near subsite -3 would allow the cellulose strand to have a flexible conformation, leading to twisting and rapid threading. Further investigation of the twisting around subsite $-4/-3$ is required.

Improved cellulase activity has a multitude of impacts on the entire biofuel production process, which comprises pretreatment, enzymatic hydrolysis, fermentation, and condensation (distillation and molecular sieving)^{15,44}, and would reduce production costs significantly. Our results demonstrate the potential of site-directed mutagenesis technology to enhance cellulase activities⁴⁵. The N62A mutant was much less productive with our *A. oryzae* host and unstable in physical processing, such as ultrafiltration, which are major drawbacks to commercial application. N62 is conserved in GH7 CBH across various microorganisms, including *T. reesei* and *P. chrysosporium* (Supplementary Fig. S1). This might suggest that this residue plays an essential role in stabilising the molecular structure. Experiments to verify whether replacing the corresponding asparagine residues with alanines is effective in GH7 CBHs from other microorganisms are warranted. Furthermore, a saturation mutagenesis method could detect another amino acid that gives a similarly increased activity to N62A without the commercial drawbacks. These findings would yield valuable information and provide further systematic analysis of the numerous diverse biorefinery applications.

Methods

Construction of *TcCel7A* encoding plasmids and transformation. An Applied Biosystems thermal cycler and KOD Plus polymerase (Toyobo) were used for all PCR experiments. The PCR reaction mixture and conditions for the polymerase were as described in the user manual. The original gene was amplified from the genomic DNA of *T. cellulolyticus* strain H1⁸ by PCR (sequence: GenBank Accession Number E39854). The expression vector was constructed by integrating the *enoA* d4 promoter⁴⁶, *agdA* terminator, and *pyrG* genes from *A. oryzae* on pMD20 (Takara). The amplified *TcCel7A* gene was inserted between the promoter and terminator regions. For alanine scanning experiments⁴⁷, point-mutated genes were prepared using an In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) with synthetic PCR primers possessing site-directed mutated bases (nine mutants: N62A, W63A, W65A, S72A, N74A, N125A, K203A, N222A, and T223A). *Aspergillus oryzae* strain NS4DLP, which lacks the *pyrG* gene and, therefore, requires uridine in the culture medium for growth, was used as the enzyme production host^{48–50}. Transformation was performed as described previously⁴⁶. *A. oryzae* cultures were maintained using medium containing 2% (w/v) dextrin hydrate, 1% (w/v) peptone, 0.1% (w/v) casamino

acid, 0.5% (w/v) KH_2PO_4 , 0.1% (w/v) NaNO_3 , and 0.05% (w/v) $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$. The pH was adjusted to 6.0 using NaOH.

Enzyme production. The medium used contained 10% (w/v) molasses, 2.5% (w/v) peptone, 2.4% (w/v) KH_2PO_4 , and 0.12% (w/v) MgSO_4 , with the pH adjusted to 6.0 using NaOH. *A. oryzae* transformants were grown in the culture medium (10 mL) at 30 °C for 3 days. The culture broth was filtered through a 0.22- μm Durapore membrane (Millipore). Production of the target enzymes contained in the filtrates was assayed by SDS-PAGE. Precast gels, and preformed running and CBB staining buffer solutions, were obtained from Atto Corporation. Prestained protein standard was purchased from Bio-Rad. Quantification of corresponding bands on the SDS-PAGE gel image was performed using a Bio-Rad imager to estimate the target enzyme concentrations⁵¹. Bovine serum albumin (BSA, Nacalai Tesque) was employed as the calibration standard. Pseudo-colour was applied using Image Lab (Bio-Rad). Calibration for the quantification of each target protein was conducted using the same software. For N62A mutant (Culture 3), see Supplementary Fig. S8. Production of WT and the four mutants, N62A, W63A, W65A, and K203A, was examined by western blotting using a specific rabbit polyclonal WT antibody as the primary antibody and the ECL prime kit including a horseradish peroxidase (HRP)-conjugated antibody (GE Healthcare). The assay was performed following manufacturer instructions. Ultrafiltration of the expressed proteins was performed using a commercially available centrifugal concentrator (MWCO: 3,000; Vivaspin, Sartorius) at 4 °C following manufacturer instructions. Deglycosylation of recombinantly produced *TcCel7A* was performed by reacting PNGase F (New England BioLabs) with the culture supernatant sample according to manufacturer instructions.

Biomass pretreatment. Corn stover pretreated with dilute acid was used to assess the activity of WT and alanine-substituted mutants^{52–54}. Dry corn stover was ground, passed through a 3-mm mesh, and mixed with twice its weight of aqueous sulfuric acid (2.5%, w/v). The mixture was kept at 150 °C for 10 min and then neutralised with NaOH. The pretreated biomass was then dried at 50 °C for 1 day and kept dry (water content, 7%) in a desiccator until being used in downstream activity assays.

Activity assay. Hydrolytic assays were carried out by incubating mixed solutions of each *A. oryzae* transformant (20 μL), 4% (w/v) reagent cellulose (100 μL ; Avicel PH-101 from Sigma or microcrystalline cellulose from Merck, cat. no. 102331) or pretreated corn stover, and 0.2 M acetate buffer (pH 5.5; 60 μL) at 50 °C for 1 h. Controls were prepared in the same way, but without substrate solids. Hydrolysis experiments were performed in duplicate. Aliquots after incubation were centrifuged (5,000 \times g), and the supernatants were assayed for cellobiose by HPLC using an HPX87 P column (Bio-rad) and RI detector (Waters)⁵⁵. Areas were calculated for the retention time range 7.8–8.5 min. Differences ascribed to the action of the target enzyme were obtained by subtracting the results without substrates from those with substrates (Supplementary Fig. S9). Hydrolysis activity toward Avicel was evaluated by colourimetry using dinitrosalicylic acid (DNS)^{56,57}. DNS solution was prepared from two solutions, $\text{KNaC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$ (Rochelle salt; 300 g) in 2 M NaOH aqueous solution (200 mL) and DNS (10 g) in purified water (400 mL; Milli-Q), which were mixed and diluted with purified water to a final volume of 1 L. To assay the reducing sugar concentration, we added the same amount of DNS solution to aliquots from reaction mixtures before and after incubation, heated these solutions at 100 °C for 5 min, and then measured their absorbance at 540 nm. We obtained the differences ascribed to the action of the target enzyme by subtracting the results without substrates from those with substrates. Hydrolysis activities toward CMC and pNPG were assayed according to the protocols of Kansarn⁵⁸, except that DNS colourimetry was employed to evaluate CMC saccharification. CMC (cat. no. C5678) and pNPG (N1627) reagents were purchased from Sigma.

Homology modelling. The primary sequence of the catalytic domain of *TcCel7A* was modified by trimming the N-terminal signal peptide (1–25 aa) and the C-terminus (461–529 aa) for modelling (Supplementary Fig. S1). A structural model for the catalytic domain was then built using *TrCel7A* (PDB:8CEL²⁰) as a template using the Discovery Studio 4.0 (Accelrys) software package. During energy minimisation to refine the positions of all heavy atoms in the protein, the structure of the substrate glycan in the template *TrCel7A* was employed and fixed. The N62A mutant was created by replacing the asparagine residue with alanine *in silico*. Protonation states at pH 7.0 were determined using the PDB2PQR server⁵⁹, and nine disulfide bonds were assigned according to the template molecules. RMSD between the homology model and other GH7 structures were calculated using the RMSD Calculator plugin in VMD⁶⁰.

MD simulation. To simulate initial threading of the cellulose substrate for WT and the N62A and N222A mutants of *TcCel7A*, the reducing end glucose (head) of the cellulose strand (cellononaose) was initially positioned at subsite –5, and other parts were created automatically (“threading” case). The systems were fully solvated with explicit solvent, and 18 Na^+ counterions were added to obtain electrostatic neutrality. All simulations were performed using AMBER14 MD software⁶¹. We employed the AMBER ff03 force field for proteins, the GLYCAM 06 force field for the cellulose strand, and the TIP3P model for water molecules. The systems were energetically minimised for the 300 steepest descent steps and equilibrated for 1 ns, with gradually reducing restraints. Finally, two 160-ns production runs were performed with different initial velocities for each system using a protocol similar to that described by Nakamura *et al.*³². Trajectory analysis was conducted using the AMBER module cpptraj and snapshot structures were visualised with VMD. An additional MD simulation was conducted for WT, in which a cellulose substrate (cellononaose) was situated to occupy all subsites (from –7 to +2) initially (“occupied” case). Force fields and protocols were the same as described above, and the production run was performed for 25 ns. Binding free energies between *TcCel7A* and the glucosyl units positioned at each subsite were calculated using the MM/GBSA method⁶², which was integrated in the AMBER14 software.

References

- Ahuja, M. R. Next generation plant biotechnology in *Biotechnology and biodiversity* (eds. Ahuja, M. R. & Ramawat, K. G.) 77–100 (Springer, 2014).
- Hadar, Y. Sources for lignocellulosic raw materials for the production of ethanol in *Lignocellulose conversion* (ed. Faraco, V.) 21–38 (Springer, 2013).
- Seay, J. R. & You, F. Biomass supply, demand, and markets in *Biomass supply chains for bioenergy and biorefining* (eds. Holm-Nielsen, J., Ehimen, E. A.) 85–100 (Elsevier, 2016).
- Nigam, P. S. & Singh, A. Production of liquid biofuels from renewable resources. *Prog. Energy Combust. Sci.* **37**(1), 52–68 (2011).
- Das Neves, M. A., Kimura, T., Shimizu, N. & Nakajima, M. State of the art and future trends of bioethanol production. *Dyn. Biochem. Process Biotechnol. Mol. Biol.* **1**(1), 1–14 (2007).
- Humbird, D. *et al.* Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute-acid pretreatment and enzymatic hydrolysis of corn stover (No. NREL/TP-5100-47764). National Renewable Energy Laboratory (NREL), Golden, CO. (2011).
- McAloon, A., Taylor, F., Yee, W., Ibsen, K. & Wooley, R. Determining the cost of producing ethanol from corn starch and lignocellulosic feedstocks. National Renewable Energy Laboratory Report (2000).
- Fukuura, M., Mitsuzawa, S., Takeda, M., Ara, T. & Shibata, D. US Patent Application 2015/0044728 (2015).
- Fukuura, M., Ara, T., Takeda, M., Shibata, D. & Mitsuzawa, S. A new approach to construct a synthetic mixture of saccharifying enzymes in *Lignocellulose degradation and biorefinery* (eds. Sakka, K. *et al.*) 262–264 (Uni, 2015b).
- Yamanobe, T., Mitsuishi, Y. & Takasaki, Y. Isolation of a cellulolytic enzyme producing microorganism, culture conditions and some properties of the enzymes. *Agric. Biol. Chem.* **51**(1), 65–74 (1987).
- Fujii, T., Hoshino, T., Inoue, H. & Yano, S. Taxonomic revision of the cellulose-degrading fungus *Acremonium cellulolyticus* nomen nudum to *Talaromyces* based on phylogenetic analysis. *FEMS Microbiol. Lett.* **351**(1), 32–41 (2014).
- Fujii, T., Koike, H., Sawayama, S., Yano, S. & Inoue, H. Draft genome sequence of *Talaromyces cellulolyticus* strain Y-94, a source of lignocellulosic biomass-degrading enzymes. *Genome Announc.* **3**(1), e00014–15 (2015).
- Gusakov, A. V. Alternatives to *Trichoderma reesei* in biofuel production. *Trends Biotechnol.* **29**(9), 419–425 (2011).
- Shinkawa, S., Mitsuzawa, S. & Tanaka, M. US Patent Application 2015/0203832 (2015).
- Mitsuzawa, S., Fukuura, M. & Shinkawa, S. Research on production of cellulosic ethanol—Creation of novel saccharifying enzyme. *Honda R&D Technical Review* **28**(2), 117–122 (2016b).
- Zhang, Y. H. P. & Lynd, L. R. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol. Bioeng.* **88**(7), 797–824 (2004).
- Imamura, C., Takahashi, H., Nakano, H., Sugihara, A. & Matsuda, H. Japan Patent Application P2008-214714 (2008).
- Nakano, H., Kobayashi, K., Ohuchi, S., Sekiguchi, S. & Yamane, T. Single-step single-molecule PCR of DNA with a homo-priming sequence using a single primer and hot-startable DNA polymerase. *J. Biosci. Bioeng.* **90**(4), 456–458 (2000).
- Miyazaki-Imamura, C. *et al.* Improvement of H₂O₂ stability of manganese peroxidase by combinatorial mutagenesis and high-throughput screening using *in vitro* expression with protein disulfide isomerase. *Protein Eng.* **16**(6), 423–428 (2003).
- Mosier, N. S., Hall, P., Ladisch, C. M. & Ladisch, M. R. Reaction kinetics, molecular action, and mechanisms of cellulolytic proteins in *Recent progress in bioconversion of lignocellulosics* (eds. Tsao, G. T. *et al.*) 23–40 (Springer, 1999).
- Lynd, L. R., Weimer, P. J., Van Zyl, W. H. & Pretorius, I. S. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**(3), 506–577 (2002).
- Breyer, W. A. & Matthews, B. W. A structural basis for processivity. *Protein Sci.* **10**(9), 1699–1711 (2001).
- Horn, S. J., Sørli, M., Vårum, K. M., Våljamæ, P. & Eijsink, V. G. Measuring processivity. *Methods Enzymol.* **510**, 69–95 (2011).
- Beckham, G. T., Bomble, Y. J., Bayer, E. A., Himmel, M. E. & Crowley, M. F. Applications of computational science for understanding enzymatic deconstruction of cellulose. *Curr. Opin. Biotechnol.* **22**(2), 231–238 (2011).
- Shang, B. Z., Chang, R. & Chu, J. W. Systems-level modeling with molecular resolution elucidates the rate-limiting mechanisms of cellulose decomposition by cellobiohydrolases. *J. Biol. Chem.* **288**(40), 29081–29089 (2013).
- Nakamura, A. *et al.* Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose. *J. Am. Chem. Soc.* **136**(12), 4584–4592 (2014).
- Divne, C., Ståhlberg, J., Teeri, T. T. & Jones, T. A. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *J. Mol. Biol.* **275**(2), 309–325 (1998).
- Payne, C. M. *et al.* Glycoside hydrolase processivity is directly related to oligosaccharide binding free energy. *J. Am. Chem. Soc.* **135**(50), 18831–18839 (2013).
- Colussi, F. *et al.* Probing substrate interactions in the active tunnel of a catalytically deficient cellobiohydrolase (Cel7). *J. Biol. Chem.* **290**(4), 2444–2454 (2015).
- Ståhlberg, J. *et al.* Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from *Trichoderma reesei*. *J. Mol. Biol.* **264**(2), 337–349 (1996).
- Bu, L. *et al.* Probing carbohydrate product expulsion from a processive cellulase with multiple absolute binding free energy methods. *J. Biol. Chem.* **286**(20), 18161–18169 (2011).
- Nakamura, A. *et al.* The tryptophan residue at the active site tunnel entrance of *Trichoderma reesei* cellobiohydrolase Cel7A is important for initiation of degradation of crystalline cellulose. *J. Biol. Chem.* **288**(19), 13503–13510 (2013).
- Kari, J. *et al.* Kinetics of cellobiohydrolase (Cel7A) variants with lowered substrate affinity. *J. Biol. Chem.* **289**(47), 32459–32468 (2014).
- Horn, S. J. *et al.* Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc. Natl. Acad. Sci.* **103**(48), 18089–18094 (2006).
- Zakariassen, H. *et al.* Aromatic residues in the catalytic center of chitinase A from *Serratia marcescens* affect processivity, enzyme activity, and biomass converting efficiency. *J. Biol. Chem.* **284**(16), 10610–10617 (2009).
- Knott, B. C., Crowley, M. F., Himmel, M. E., Ståhlberg, J. & Beckham, G. T. Carbohydrate–protein interactions that drive processive polysaccharide translocation in enzymes revealed from a computational study of cellobiohydrolase processivity. *J. Am. Chem. Soc.* **136**(24), 8810–8819 (2014).
- Kostylev, M. *et al.* Cel48A from *Thermobifida fusca*: structure and site directed mutagenesis of key residues. *Biotech. Bioeng.* **111**(4), 664–673 (2014).
- Mitsuzawa, S., Kimura, K., Tanaka, M. & Shinkawa, S. US Patent Application 2016/0160200 (2016).
- Christensen, T. *et al.* High level expression of recombinant genes in *Aspergillus oryzae*. *Nat. Biotech.* **6**(12), 1419–1422 (1988).
- Shibuya, I., Tsuchiya, K., Tamura, G., Ishikawa, T. & Hara, S. Overproduction of an α -amylase/glucoamylase fusion protein in *Aspergillus oryzae* using a high expression vector. *Biosci. Biotech. Biochem.* **56**(10), 1674–1675 (1992).
- Igarashi, K., Wada, M., Hori, R. & Samejima, M. Surface density of cellobiohydrolase on crystalline celluloses. *FEBS J.* **273**(13), 2869–2878 (2006).
- Momeni, M. H. *et al.* Structural, biochemical, and computational characterization of the glycoside hydrolase family 7 cellobiohydrolase of the tree-killing fungus *Heterobasidion irregulare*. *J. Biol. Chem.* **288**(8), 5861–5872 (2013).
- Pace, C. N. *et al.* Contribution of hydrogen bonds to protein stability. *Protein Sci.* **23**(5), 652–661 (2014).
- Tsuchida, Y. & Tsukagoshi, N. Research on production of cellulosic ethanol—Creation of novel saccharifying enzyme. *Honda R&D technical review* **28**(2), 108–116 (2016).

45. Molina-Espeja, P. *et al.* Beyond the outer limits of nature by directed evolution. *Biotechnol. Adv.* **34**(5), 754–767 (2016).
46. Toda, T. *et al.* Deletion analysis of the enolase gene (*enoA*) promoter from the filamentous fungus *Aspergillus oryzae*. *Curr. Genet.* **40**(4), 260–267 (2001).
47. Cunningham, B. C. & Wells, J. A. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244**(4908), 1081–1085 (1989).
48. Marui, J. *et al.* Penicillin biosynthesis in *Aspergillus oryzae* and its overproduction by genetic engineering. *J. Biosci. Bioeng.* **110**(1), 8–11 (2010).
49. Mizutani, O. *et al.* A defect of LigD (human Lig4 homolog) for nonhomologous end joining significantly improves efficiency of gene-targeting in *Aspergillus oryzae*. *Fungal. Genet. Biol.* **45**(6), 878–889 (2008).
50. Yamada, O., Lee, B. R. & Gomi, K. Transformation system for *Aspergillus oryzae* with double auxotrophic mutations, *niaD* and *sC*. *Biosci. Biotechnol. Biochem.* **61**(8), 1367–1369 (1997).
51. Miles, A. P. & Saul, A. Using SDS-PAGE and scanning laser densitometry to measure yield and degradation of proteins in *The Protein Protocols Handbook* (ed. Walker, J. M.) 487–496 (Springer, 2009).
52. Mosier, N. *et al.* Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* **96**(6), 673–686 (2005).
53. Hendriks, A. T. W. M. & Zeeman, G. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresour. Technol.* **100**(1), 10–18 (2009).
54. Taherzadeh, M. J. & Karimi, K. Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review. *Int. J. Mol. Sci.* **9**(9), 1621–1651 (2008).
55. Resch, M. G., Baker, J. O. & Decker, S. R. Low solids enzymatic saccharification of lignocellulosic biomass. Laboratory Analytical Procedure/Technical Report NREL/TP-5100-63351 prepared for National Renewable Energy Laboratory: Golden, CO (2015).
56. Miller, G. L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31**(3), 426–428 (1959).
57. Ghose, T. K. Measurement of cellulase activities. *Pure Appl. Chem.* **59**(2), 257–268 (1987).
58. Kansarn, S., A novel concept for the enzymatic degradation of native cellulose by *Acremonium cellulolyticus*. Thesis, Shizuoka University (2000).
59. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**(suppl_2), W665–W667 (2004).
60. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual MDynamics. *J. Mol. Graph.* **14**(1), 33–38 (1996).
61. Case, D. A. *et al.* AMBER 14, University of California, San Francisco (2014).
62. Massova, I. & Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discov.* **18**(1), 113–135 (2000).

Acknowledgements

We thank Katsuya Gomi (Tohoku University) and Masayuki Machida (The National Institute of Advanced Industrial Science and Technology) for technical support and assistance with the transformation of *Aspergillus oryzae*. We thank Simon Partridge, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Author Contributions

S.M., M.F., S.S., and K.K. designed and performed the experiments. S.M. and T.F. designed and performed the simulations. S.M. and T.F. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-16434-x>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017