

**Molecular and functional analysis of  
glycoside hydrolase family 7 cellobiohydrolases from filamentous fungi**

**糖質加水分解酵素ファミリー7に属する糸状菌由来セロビオヒドロラーゼの分子機能解析**

**Akihiko Nakamura**

**中村 彰彦**

# Contents

<b>1. Introduction</b>	<b>7</b>
1.1 Cellulose	
1.1.1 Definition of cellulose	8
1.1.2 Crystal polymorphs of cellulose	9
1.2 Cellulases	
1.2.1 Research history of cellulases and cellulose degradation model of fungi	12
1.2.2 Classification and general characterization of cellulase	
1.2.2.1 Enzyme commission number	16
1.2.2.2 Attack point of cellulose and processivity	16
1.2.2.3 Reaction mechanisms	17
1.2.2.4 Glycoside hydrolase family	18
1.3 GH family 7 Cellobiohydrolases	
1.3.1 General characteristics of GH7 cellobiohydrolases	21
1.3.2 Amino acid sequence comparison between CBH and EG	22
1.3.3 Comparison of loop structures	
between CBH form ascomycete and basidiomycete	24
1.3.4 Recent advances in GH7 cellobiohydrolase research	26
1.4 Aim of this thesis	32
<b>2. Materials and methods</b>	<b>33</b>
2.1 Enzyme production	
2.1.1 <i>TrCel7A</i> mutants	34
2.1.2 <i>PcCel7C</i> and <i>PcCel7D</i>	34

2.2	Enzyme purification	
2.2.1	<i>TrCel7A</i> and <i>TrCel7B</i> and <i>TrCel7A</i> mutants	35
2.2.2	<i>PcCel7C</i> and <i>PcCel7D</i>	36
2.3	Cellulose preparation	
2.3.1	Crystalline cellulose	37
2.3.2	Amorphous cellulose	38
2.4	Enzyme characterization	
2.4.1	Activity comparison to cellulose	39
2.4.2	Determination of kinetic parameters to <i>pNP</i> -cellobioside	39
2.4.3	Activity comparison to cello-oligosaccharides	40
2.4.4	Comparison of kinetic parameters to cellotetraose and cellopentaose	40
2.4.5	Molecular dynamics simulation of catalytic domains of WT and W40A	41
2.4.61	High-speed AFM observation	41
<b>3.</b>	<b>Results and discussion</b>	<b>43</b>
3.1	Analysis of cellulose recognition by CBH unique tryptophan residues	
3.1.1	Comparison of kinetic parameters to <i>pNP</i> -cellobioside	44
3.1.2	Comparison of kinetic parameters to cellotetraose	47
3.1.3	Activity comparison to crystalline cellulose	51
3.1.4	Estimation of the role of Trp40 residue by molecular dynamics simulation	52
3.2	Analysis of processive reaction of GH7 cellobiohydrolases	
3.2.1	General characterization of GH7 CBH from ascomycete and basidiomycete	54
3.2.2	High-speed AFM observation	57
3.2.3	Validation of processivity	63

<b>4. Conclusion</b>	65
<b>References</b>	68
<b>Acknowledgements</b>	77
<b>List of publications</b>	78
<b>Appendix</b>	79

## Abbreviations

AA	auxiliary activity
BMCC	bacterial microcrystalline cellulose
CAZy	carbohydrate active enzymes
CBD	carbohydrate binding domain
CBH	cellobiohydrolase
CBM	carbohydrate binding module
CD	catalytic domain
CMC	carboxy-methyl cellulose
DO	dissolved oxygen
EC	enzyme commission
EG	endoglucanase
GH	glycoside hydrolase
HCC	highly crystalline cellulose
HOPG	highly oriented pyrolytic graphite
HPLC	high performance liquid chromatography
HS-AFM	high speed atomic force microscopy
PDB	protein data bank
<i>p</i> NPC	<i>p</i> -nitrophenyl- $\beta$ -D-cellobioside
<i>p</i> NPL	<i>p</i> -nitrophenyl- $\beta$ -D-lactoside
PVDF	polyvinylidene difluoride
RFP	red-fluorescent protein
Trp	tryptophan
WT	wild type



# 1. Introduction

## 1.1 Cellulose

### 1.1.1 Definition of cellulose

Cellulose is the main component of plant cell walls and the most abundant biological resource on Earth. Cellulose fibril structure is constructed from cellulose molecule chains. Single cellulose chain is constructed by D-glucopyranose residues which are linked by  $\beta$ -1,4 glycosidic bonds without any branches (Fig. 1-1). Neighboring glucose units are  $180^\circ$  inverted with respect to each other and therefore cellobiose (4-O-( $\beta$ -D-glucopyranosyl)- $\beta$ -D-glucopyranose) is the repeating units. The chain has a reducing end with free C-1 hydroxyl and a non-reducing end, in which the C-1 hydroxyl is linked to the C-4 of the next glucose residue.

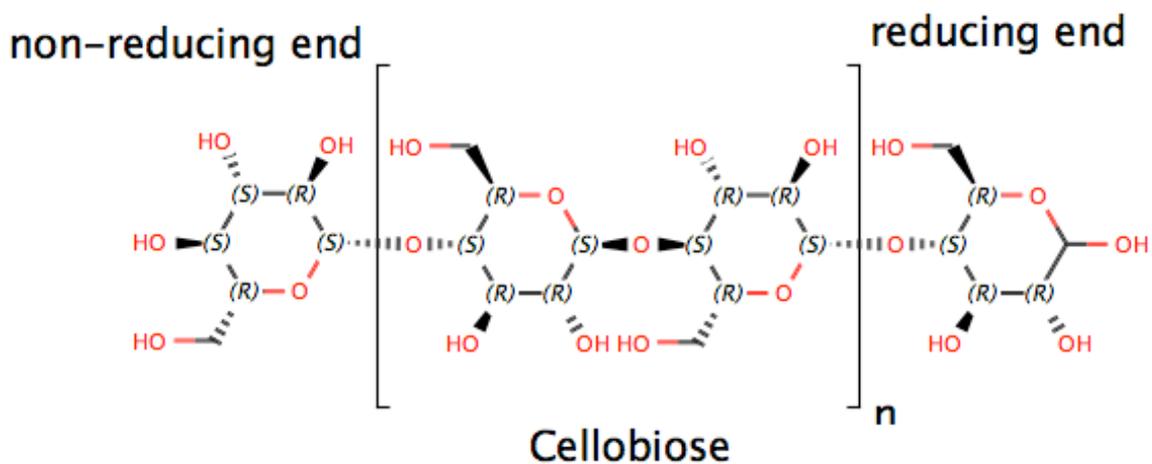


Fig. 1-1 structure of cellulose chain

Cellulose has two different high-order structures, amorphous and crystal. Crystalline cellulose is a highly ordered structure of cellulose molecule chains, which is stabilized by inter-chain and intra-chain hydrogen bonds and hydrophobic interactions. Cellulose microfibrils can be assembled in some different crystal forms. On the other hand, amorphous cellulose is a non-ordered structure and that is considered as an aggregation of cellulose chains. Because of the randomness of the structure, the exact structure of amorphous cellulose have not been determined. The clear discrimination is not crystal.

### 1.1.2 Crystal polymorphs of cellulose

Crystal polymorphs of cellulose and the relationships between each form are shown in Fig. 1-2. Native cellulose is a mixture of cellulose  $I_\alpha$  and  $I_\beta$  (1). The cellulose from bacteria or alga are basically  $I_\alpha$ -rich (2,3) and that from plant or tunicates are  $I_\beta$ -rich (4,5). Cellulose  $I_\alpha$  irreversibly transforms to cellulose  $I_\beta$  at 260°C in 0.1 M NaOH (6). Crystal sizes and shapes are also different depending on the origin. Cellulose microfibrils from the alga *Valonia macrophysa* has been shown to be constructed of about 1200 cellulose chains (7) and plant cellulose microfibrils contain about 24 chains (8). These differences are thought to be related to the arrangement of complexes of cellulose synthase enzymes (8-11). Cellulose II was irreversibly prepared from cellulose I by regeneration or mercerization. Regeneration means a re-crystallization of cellulose in a poor solvent after a dissolution in the appropriate solvent, like a mixture solution of copper and ethylene diamine. Mercerization is a swelling and re-arrangement of chains in the crystal in a high concentration of sodium hydroxide solution (12,13). Cellulose  $III_I$  and  $III_{II}$  are made from cellulose I and Cellulose II respectively. The complex of cellulose and ammonia is formed

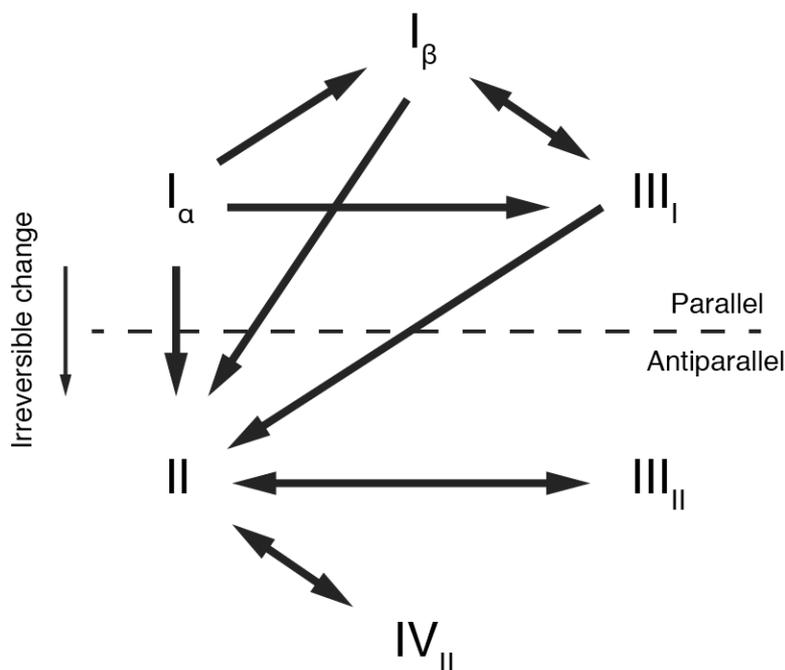


Fig. 1-2 Inter-conversion scheme among polymorphous of cellulose

by the treatment of liquid ammonia or supercritical ammonia, and then cellulose III is prepared by removal of ammonia molecules from the complex (14,15). Ethylenediamine also has been shown to make a complex with cellulose and transform crystal to cellulose III (16). Cellulose III<sub>I</sub> and III<sub>II</sub> can be transformed back to cellulose crystal I and II by hydrothermal treatment (17). Cellulose IV<sub>II</sub> was prepared at 150°C by de-acetylation of cellulose triacetate in a mixture of methylamine, dimethyl sulfoxide, and water (18). In contrast, Cellulose IV has not been obtained from cellulose I (19).

Crystal structures of cellulose have been studied by X-ray and neutron diffraction analysis to determine the structure of main chains and hydrogen bond network among the chains (Fig. 1-3 and 1-4). Cellulose I<sub>α</sub> is formed by only one conformation of parallel cellulose chains, and stabilized by two strong hydrogen bonds (O3-H to O5 and O2-H to O6) in cellulose chain and one hydrogen bond (O6-H to O3) between neighboring chains (20). Molecular sheets slide directionally each other per a quarter of c axis length (about half of glucose length). Crystal structure of cellulose I<sub>β</sub> is similar to I<sub>α</sub>, but I<sub>β</sub> is formed by two different conformations of parallel cellulose chains (21). The two conformations of chains are named as 'Origin' and 'Center' due to the position at the unit cell. Cellulose I<sub>β</sub> has also two intra-chain hydrogen bonds and one hydrogen bond between chains. Molecular sheets are positioned zig-zag each other per a quarter of c axis length. In the I<sub>α</sub> and I<sub>β</sub>, there is a multi possibility of hydrogen bonds because the position of O6 proton has multiplicity. Molecular Sheets of them are packed by only hydrophobic interaction. Cellulose II is formed by two conformations of anti-parallel chains (origin and center), and stabilized by a hydrogen bond (O3-H to O5 or O3-H to O6) in chain and two hydrogen bonds (O6-H to O2 and O2-H to O6) between chains (22,23). The sheets are zig-zag forms, and two hydrogen bonds (O6-H to O6 and O2-H to O2) connect the sheets. The sheets slide also zig-zag each other per a quarter of c axis length. Cellulose III<sub>I</sub> was formed by single conformation of parallel cellulose chains, and stabilized by a hydrogen

bond (O3-H to O5 or O3-H to O6) in chain and two hydrogen bonds (O2-H to O6 and O6-H to O2) between neighboring chains or zig-zag sheets (24).

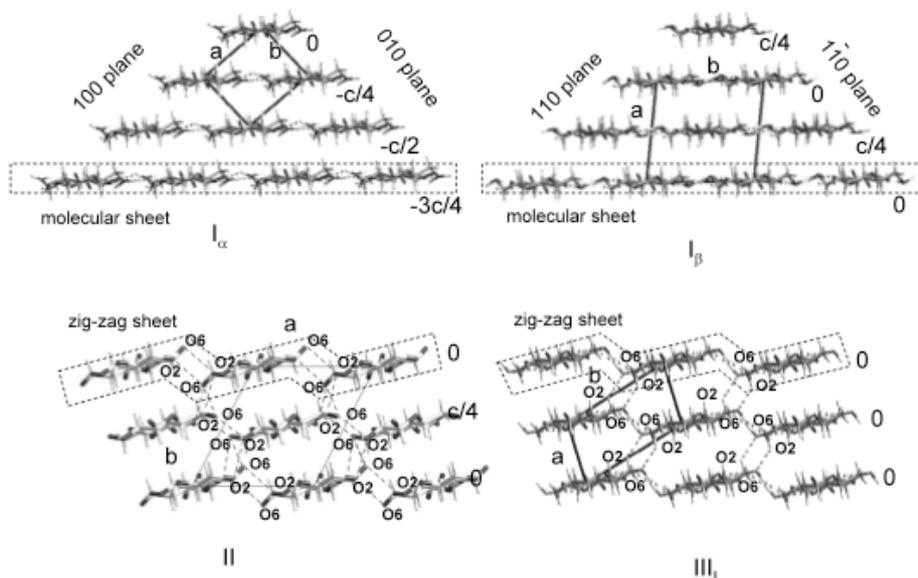


Fig. 1-3 Comparison of chain packing and hydrogen bonds among the cellulose polymorphs (from ref (25)).

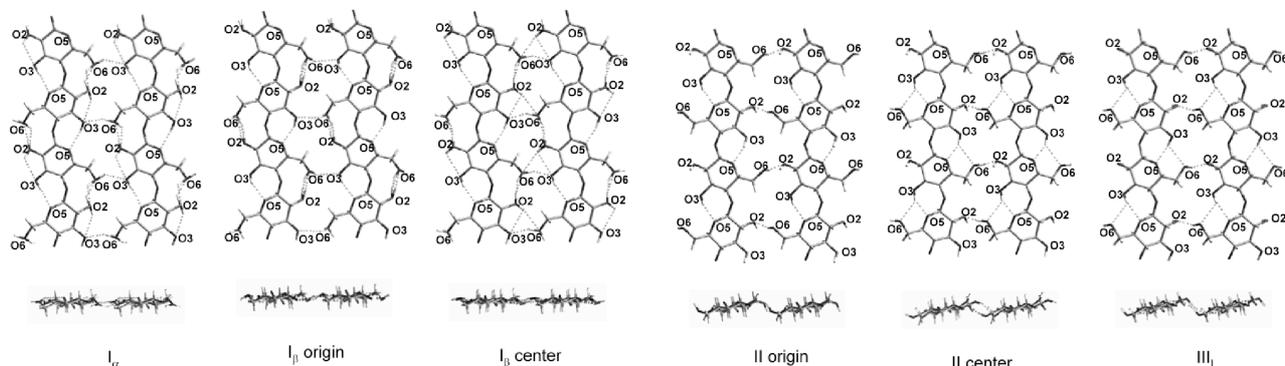


Fig. 1-4 Hydrogen bonds between the neighboring chain (from ref (25)).

## 1.2 Cellulases

### 1.2.1 Research history of cellulases and cellulose degradation model of fungi

The history of cellulase researches is summarized in Fig. 1-5. First cellulose degradation by microorganisms was observed in 1850, the word 'cellulase' was used in 1912, and the presence of cellulase was experimentally suggested in 1927 (26,27). The first model of enzymatic cellulose degradation is the C<sub>1</sub>-C<sub>x</sub> theory, which has been suggested in 1950 (28). In this theory, native cellulose is degraded with two systems; C<sub>1</sub> degrades cellulose to a more easily digestible form and C<sub>x</sub> hydrolyzes that to soluble short saccharides. The difference between cellulolytic organisms and noncellulolytic organisms was defined by with or without C<sub>1</sub> system. In 1954, cellulases produced by microorganisms were shown as multiple components, and their synergistic effect was reported (29). During 1960s and early half of 1970s, many studies about purifications of cellulase components have been reported to determine the enzymes act as C<sub>1</sub> or C<sub>x</sub> (30). Because C<sub>x</sub> system was detected to act on carboxy-methyl cellulose (CMC), the enzyme was called CM-cellulase. In 1965, C<sub>x</sub> components were separated to endo-glucanase (EG) and 'exo-glucanase' (means  $\beta$ -glucosidase in this paper) (31). Additionally, in 1972 and 1973, C<sub>1</sub> component was determined as cellobiohydrolase (CBH), which produces cellobiose from cotton cellulose (32,33). In 1965, the base of next model; endo-exo theory was also suggested (31). The theory mentioned that endo-glucanases cut the molecular chain of cellulose randomly and increase the chain ends for exo-glucanases, but did not fully explain. The revised theory of endo-exo synergy degradation of cellulose was suggested by K. E. Eriksson around 1970 (34,35) and experimentally supported in 1975 (36). After the suggestion of this theory, many endo-glucanases and exo-glucanases were purified and characterized (37-39). In 1980, two cellobiohydrolases (CBH I and II) were purified from crude enzyme of *Trichoderma reesei*, and exo-exo synergy was suggested (40). In this paper, the two CBHs were first identified by the N-terminal amino acid sequence. The gene of CBH I was first

cloned from *T. reesei* in 1983 (41). Additionally, from *T. reesei*, the gene of EG I was cloned in 1986, and the gene of CBH II was cloned in 1987 (42,43). After the cloning of cellulases, each enzymes have been distinguished by the gene of the enzymes and expected sequence of amino acids. Moreover, heterologous expression and characterization of fungal cellulases have been started (44). About the activity of cellulase, the difference of specificity to chain ends between CBH I and CBH II was found in 1989 via cello-oligosaccharides labeled with 4-methylumbelliferyl and reduced alkali-swollen cellulose (45,46). The reaction point of CBH I was also tested in 1992 with cello-oligosaccharides labeled with tritium (47). By the development of X-ray crystallography, X-ray structure of catalytic core of CBH II was solved in 1990, and that of CBH I was solved in 1994 (48,49). Furthermore, from *Thermomonospora fusca*, the X-ray structure of the endo-glucanase, which has similarity with CBH II has been solved in 1993 (50). By this result, the structural difference between cellobiohydrolase and endo-glucanase was clarified. The relationship between structure and enzyme mechanism of cellulases was discussed in 1995 (51). In this paper, the processivity of cellobiohydrolase was suggested due to the tunnel like structure of them. After these discoveries, the model of cellulose degradation was further revised in 1997; CBH I and CBH II were suggested to degrade cellulose from reducing-end and non-reducing end respectively, and endo-glucanases were suggested to cut the amorphous region to make the chain end for CBH reaction (52). During 1990 to 2000, the names of enzymes were changed systematically based on the family clustering. The fundamental method of family clustering was reported in 1989, and classification into families was suggested in 1991, but the family classification was widely accepted only after the online database was opened in 1998 (53,54). Recently, one of the family of endoglucanase was deleted. Previously glycoside hydrolase family 61 (GH61) enzymes were thought to be endo-glucanases, but the redox activity of them was suggested in 2010 due to their structural similarity to the CBP21s (chitin binding proteins),

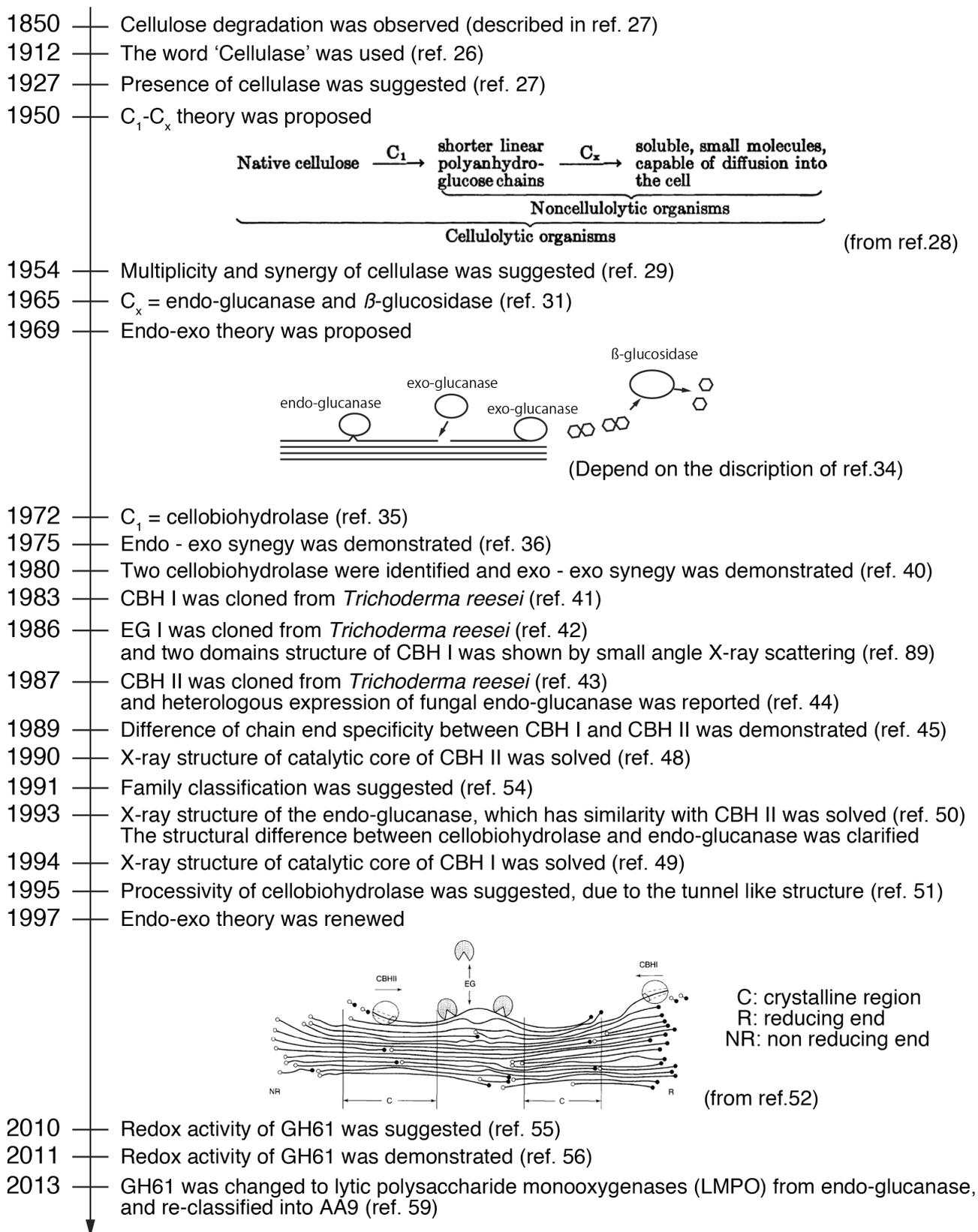


Fig. 1-5 Chronology of cellulase research

which have oxidative activity for chitin and synergy with chitinase (55). The oxidative activity of GH61 have been reported in 2011 and 2012 (56-58). The name of GH61 has been changed to lytic polysaccharide monooxygenases (LMPO), and re-classified into auxiliary activity (AA) family 9 (59).

To check the research trend of cellulase, the number of published articles about cellulase between 1990 to 2012 was analyzed (Fig. 1-6). The number of articles was stuck during 1991 to 2006, but the number was remarkably increased after 2007. When one more keyword (gene, structure or biomass) was added, the recent trend was clarified. About the gene and structure, the articles appeared after the first report of each technique, and the numbers of articles slightly increased recently. On the other hand, the number of articles including 'biomass' as a keyword slightly increased during 1991 to 2006, but drastically increased after 2007. The value in 2012 was over 300, and occupied about one third of all articles about cellulase. This result indicates that the recent research trend of cellulase is application for biomass conversion.

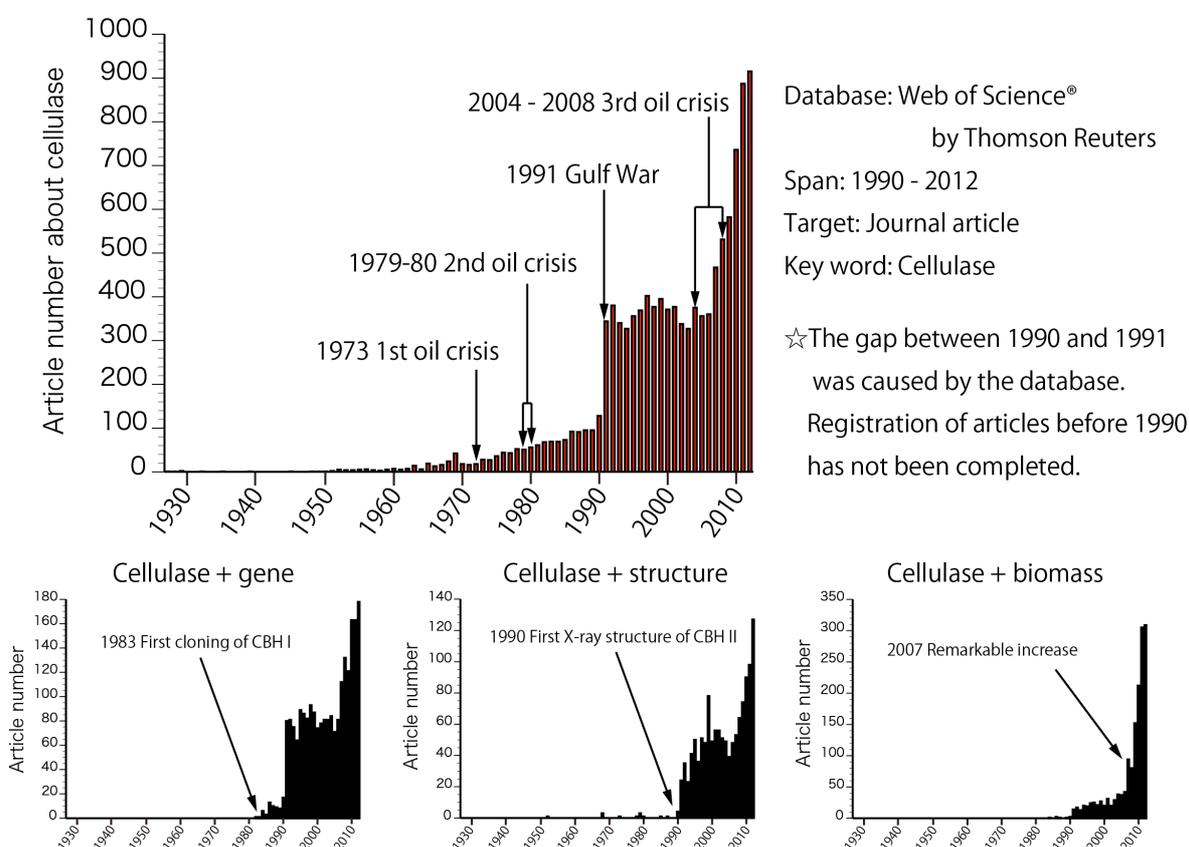


Fig. 1-6 Analysis of article number about cellulase

## 1.2.2 Classification and general characterization of cellulase

### 1.2.2.1 Enzyme commission number

Enzyme commission (EC) number represents the chemical reaction of the enzyme (Fig. 1-7). Hydrolases are assigned to EC 3 category, and cellulases are classified in three EC numbers. EC 3.2.1.4 is endo- $\beta$ -1,4 glucanase, which catalyzes endo-hydrolysis of  $\beta$ -1,4 glucosidic linkages in cellulose, lichenan and other  $\beta$ -glucans (60,61). EC 3.2.1.91 and EC 3.2.1.176 are non-reducing-end cellobiohydrolase and reducing-end cellobiohydrolase respectively, which mainly produce cellobiose from each end of cellulose (62,63).

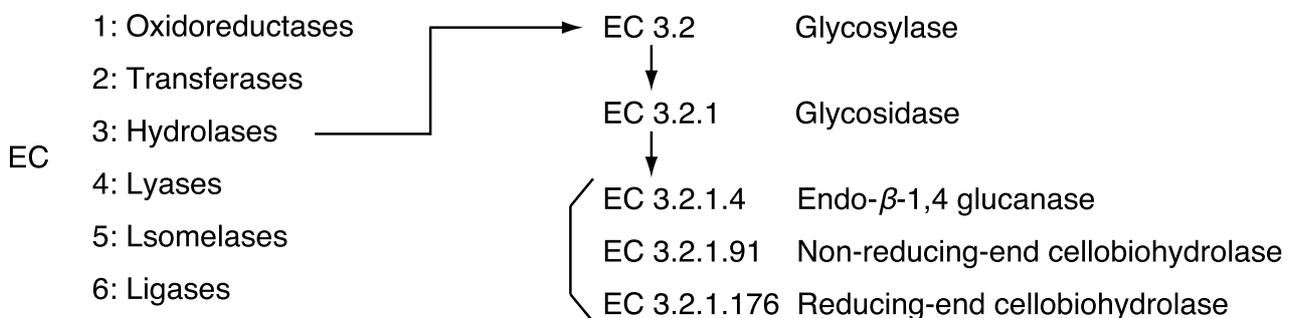


Fig. 1-7 EC numbers of cellulases

### 1.2.2.2 Attack point of cellulose and processivity

Cellulases are divided in three types by the reaction point on the cellulose (64) (Fig. 1-8). First is the endo-attack which cleaves a glycoside bond not only from the ends of a cellulose molecule chain, but also bonds within a molecular chain. Second is non-reducing-end exo-attack, which hydrolyzes cellulose from non-reducing end (including non-reducing-end CBH), and the last is reducing-end exo-attack, which hydrolyzes from

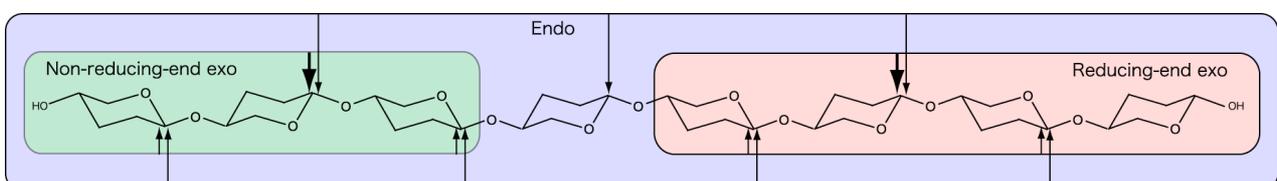


Fig. 1-8 Exo attack and Endo attack

the other end (including reducing- end CBH).

Processivity is defined as the number of reaction without dissociation after the formation of enzyme-substrate complex. This concept has been established about the DNA polymerase I and amylase (65,66). In case of cellulase, many cellobiose multiple units of cellobiose) are produced from cellulose by processive reaction (Fig. 1-9). Although processivity is independent of the substrate intake pattern, processive cellulases are sometimes divided into exo-processive and endo-processive.



Fig. 1-9 Processivity of cellulase

### 1.2.2.3 Reaction mechanisms

Cellulase is an enzyme which catalyze the hydrolysis of  $\beta$ -1,4 linkages between glucose residues (Fig. 1-10), and two different reaction mechanisms have been found in cellulases (Fig. 1-11). One is inverting mechanism, and the other is retaining mechanism. The anomeric conformations of products are converted with the inverting mechanism, but those are kept with retaining mechanism (67). In the inverting mechanism, it is said that the water molecule, which is activated by general base residue, attacks to the C1 atom of substrate and general acid residue donates the proton to O1 atom of substrate. This reaction is

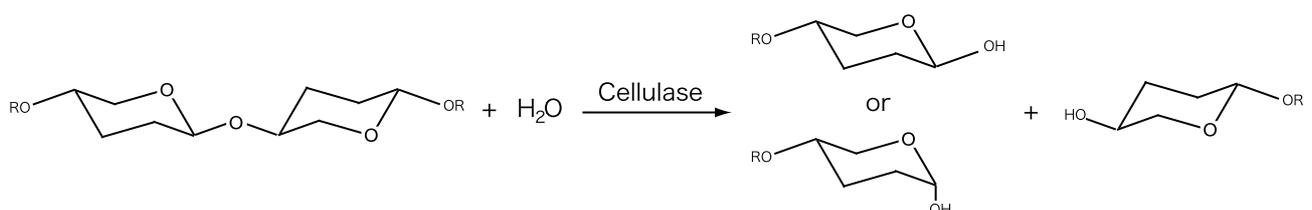
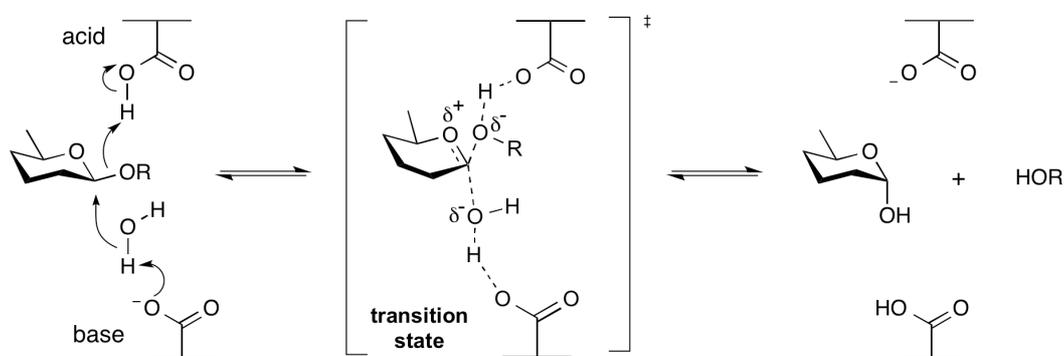


Fig. 1-10 Reaction scheme of cellulase

supposed to be catalyzed with single-displacement mechanism intermediating oxocarbenium ion-like transition states. The catalytic residues are normally acidic residues and the distance between them is longer than that of retaining enzymes (68). On the other hand, retaining reaction is achieved via two steps; glycosylation and deglycosylation steps (69). In the first step, the nucleophile residue attacks to the C1 atom, and the acid/base residue protonates the glycosidic oxygen atom. After the step, glycosyl enzyme intermediate is formed. In the second step, the intermediate is hydrolyzed by the water molecule deprotonated by the acid/base residue.

#### Inverting mechanism



#### Retaining mechanism

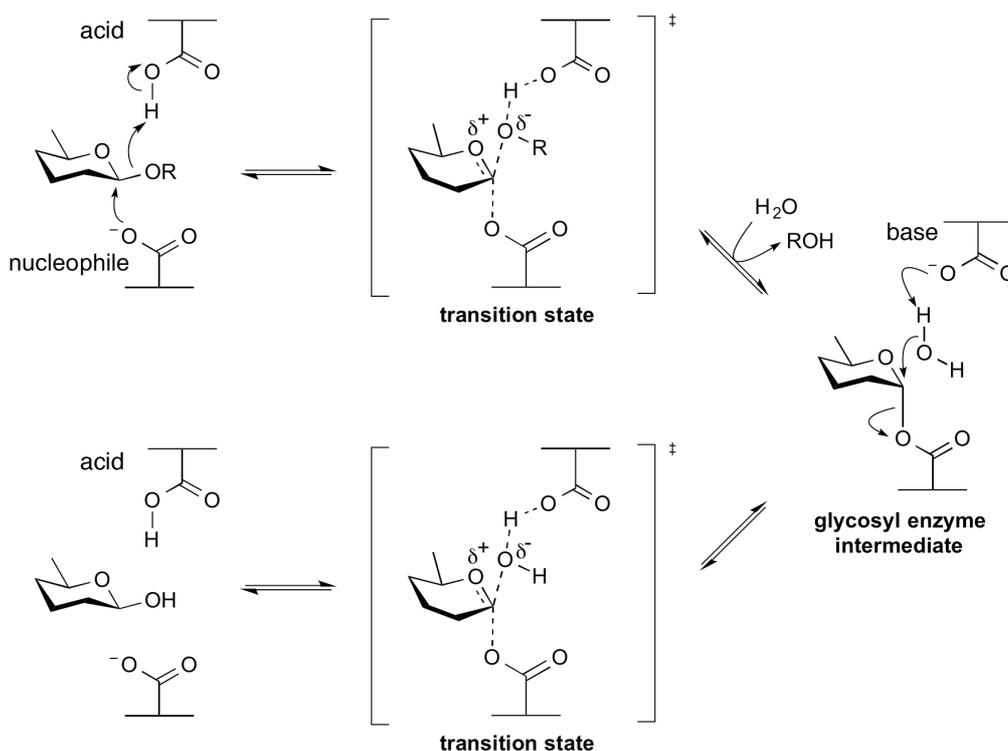


Fig. 1-11 Reaction mechanisms of cellulase (from Cazypedia and modified)

#### 1.2.2.4 Glycoside hydrolase family

Carbohydrate active enzymes and associated modules (like as carbohydrate-binding module: CBM) are classified into 'Families' based on their amino acid sequences and summarized in the CAZy database (70). Cellulases are included in several glycoside hydrolase (GH) families (53,54). At this time (Dec. 4, 2013), 133 GH families exist, and cellulases are classified in 12 families (GH5, 6, 7, 8, 9, 12, 44, 45, 48, 51, 74 and 124) (Fig. 1-12 and Table 1-1). Among them, exo activities have been reported to occur only in four families (GH 6, 7, 9, 48). Although GH61 was previously classified as a cellulase, recently it was moved to auxiliary activity (AA) family 9 due to the redox activity (56,59).

A clan is a group of GH families that share significant similarity in their 3-dimensional structure, catalytic residues and mechanism (71). 14 clans (GH-A to N) are suggested in CAZy now.

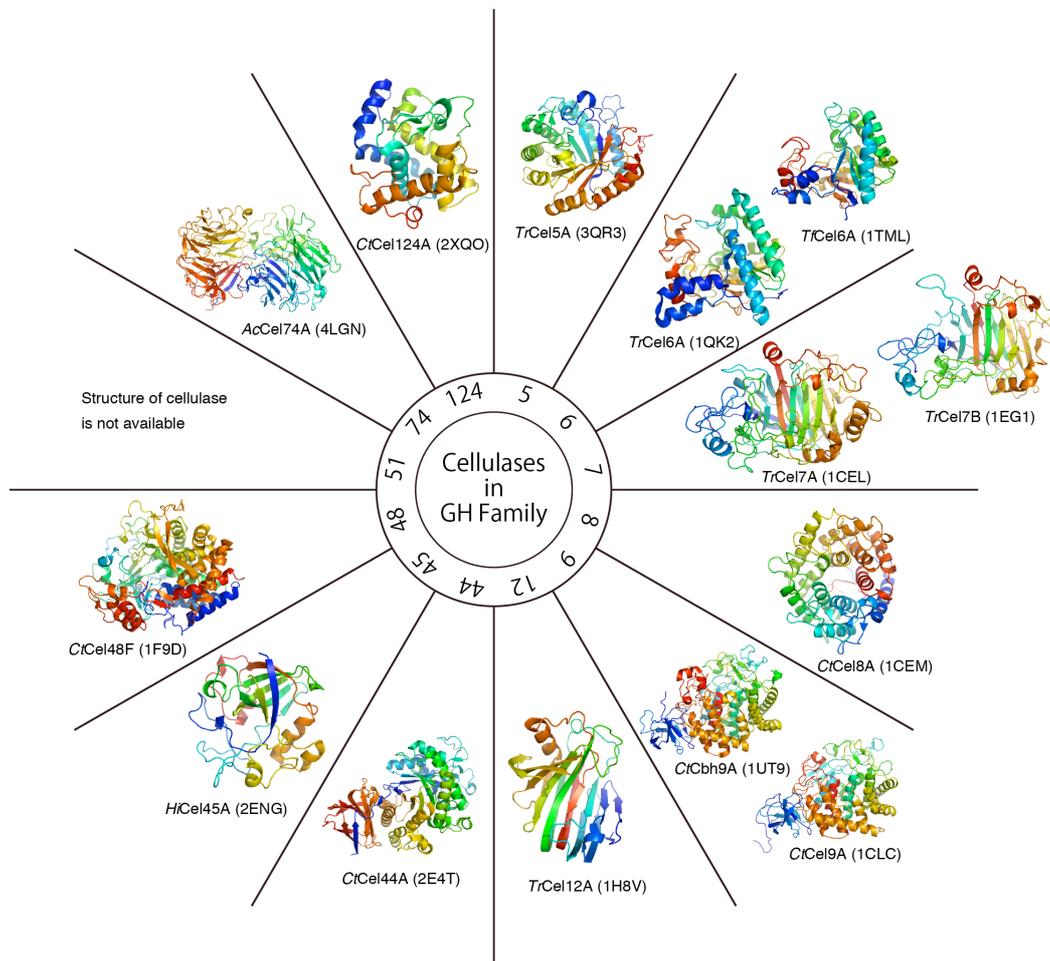


Fig. 1-12 Structures of cellulases classified in GH families

Table 1-1 Family classification of cellulase

Family	Reaction mechanism (base•acid) or (nucleophile •acid/base)	Attack point	Clan (Fold)	CBM
5	Retainig (Glu•Glu)	Endo	GH-A ( (α/β)8 barrel )	CBM1 CBM2 CBM3 CBM4 CBM6 CBM9 CBM10 CBM11 CBM17 CBM28
6	Inverting (Asp•Asp)	Endo Exo (Non-reducing end)	- ( (β/α)7 barrel )	CBM1 CBM2 CBM10
7	Retainig (Glu•Glu)	Endo Exo (Reducing end)	GH-B (β-jelly roll)	CBM1
8	Inverting (Asp•Glu)	Endo	GH-M ( (α/α)6 barrel )	-
9	Inverting (Asp•Glu)	Endo Exo (Non-reducing end)	- ( (β/α)8 barrel )	CBM2 CBM3 CBM4 CBM9 CBM10
12	Inverting (Glu•Glu)	Endo	GH-C (β-jelly roll)	CBM1 CBM2
44	Retainig (Glu•Glu)	Endo	- ( (β/α)8 barrel )	CBM3
45	Retainig (Asp•Asp)	Endo	- ( double ψ β-barrel )	CBM1 CBM2 CBM10
48	Inverting ( - •Glu)	Endo Exo (Reducing end)	GH-M ( (α/α)6 barrel )	CBM2 CBM3 CBM10
51	Retainig (Glu•Glu)	Endo	GH-A ( (α/β)8 barrel )	CBM11
74	Inverting (Asp•Asp)	Endo	- (two 7-bladed β-propeller in tandem)	CBM1
124	Inverting ( - •Glu)	Endo	- (α+β "lysozyme fold")	-

## 1.3 GH family 7 Cellobiohydrolases

### 1.3.1 General characteristics of GH7 cellobiohydrolases

GH7 cellobiohydrolase (CBH) is the major enzyme produced by filamentous fungi in cellulolytic culture (38). Interestingly, GH7 CBH and EG are unique enzymes for fungi. GH7 CBH was first purified as an exo-1,4-beta-glucanase from crude enzymes of *Trichoderma koningii* in 1972 and *Sporotrichum pulverulentum* (= *Phanerochaete chrysosporium*) in 1975 (32,72). In 1980, GH7 cellobiohydrolase from *Trichoderma reesei* was named as CBH I due to the discovery of another cellobiohydrolase: CBH II (GH6 cellobiohydrolase) (40). The gene of CBH I was cloned from *T. reesei* and its expected amino acid sequence has been available since 1983, and CBH I has been classified in GH family 7 since 1991 (41,54). The chain end specificity (GH7 CBH hydrolyze cellulose from a reducing-end) was reported in 1989, and the EC number (EC 3.2.1.176) was assigned to GH7 CBH in 2011 (45,46,63). Furthermore, the processive reaction of *TrCel7A* on the molecular chain of cellulose on the crystalline cellulose surface was first directly observed by high speed atomic force microscopy in 2009 (73).

Some GH7 CBHs have an additional domain, which is carbohydrate binding domain (CBD) classified in carbohydrate-binding module (CBM) family 1 (Fig. 1-13). These two domains, catalytic domain and CBD, are connected by an O-glycosylated linker region. Catalytic domain (CD) of GH7 CBH is constructed of about 430 amino acids, and has

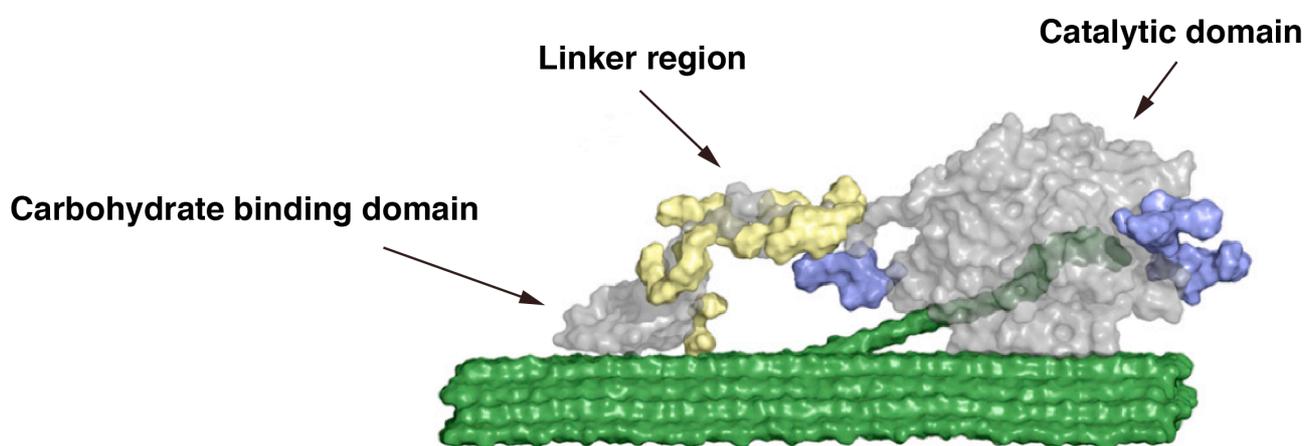


Fig. 1-13 Structure of *TrCel7A* (Modified figure of ref.(74))

$\beta$ -jelly roll fold (49). The CD of GH7 cellobiohydrolase has ten subsites (-7 to +3) and a tunnel-like structure with loop regions stabilized by disulfide bridges (75,76). The catalytic mechanism of GH7 CBH is retaining (77). GH7 CBH has the conserved motif (Glu-X-Asp-X-X-Glu) at the catalytic center, and former and later glutamic acid act as the catalytic nucleophile residue and acid/base residue respectively under assistance of aspartic acid (78-80). CBD is constructed of about 30 amino acids and also stabilized by disulfide bridges (81). Three hydrophobic amino acids form a flat binding surface and they are located per cellobiose units (82). Linker region includes many serine, threonine and proline residues and that is highly O-glycosylated (83). Recently, the effect of the linker region to binding on the cellulose surface was suggested by molecular dynamics simulation (74).

### 1.3.2 Amino acid sequence comparison between CBH and EG

*TrCel7A* is the most studied CBH classified in GH7 from the ascomycete *Trichoderma reesei*, and the X-ray structure of CD including substrate along the subsites is available (84). The structure of *TrCel7A* CD with cellononaose is illustrated in Fig. 1-14 A, and the glucose faces are shown in Fig. 1-14 B. The face, which the carbon atoms constructing glucose ring are clockwise, is named as  $\alpha$ -face, and the other face is named as  $\beta$ -face (85). Cellulose chain is twisted 90 degrees at the subsite -3 and inverted at the subsite -2. Four tryptophan residues are located along the subsites, and Trp40 residue interacts with the  $\alpha$ -face of glucose, but the other residues (Trp38, Trp367 and Trp376) interact with the  $\beta$ -face of glucose. Among them, the two tryptophan residues (Trp38 and Trp40) are located at the entrance of the tunnel-like structure, and the others are constructing the active site. Recently, a lot of information of genes coding GH7 enzymes are available by the studies about cloning of the genes and genome analysis of fungi. In addition, informations concerning thier activity is reported and summarized in the CAZy database (<http://www.cazy.org>). The alignment of GH7 enzymes, made by the MAFFT ver.7 software

(<http://mafft.cbrc.jp/alignment/server/>), is shown in Fig. 1-14. The aligned sequences are classified to in CBH-type and EG-type sequences (Fig. 1-14). Among the four tryptophan residues, the residues corresponding to Trp367 and Trp376 in *TrCel7A* are conserved in all GH7 enzymes, but the residues corresponding to Trp40 and Trp38 are unique for CBH type enzymes. Additionally, the residues conserved only in CBH were marked in Fig. 1-15. These residues are all constructing subsite from -7 to -3, which is near the entrance of the tunnel (Fig. 1-15).

Interestingly, the tryptophan residue at the entrance of the tunnel is also conserved among the GH family 6 CBHs, and it has been reported that the enzymes lose their activity for crystalline cellulose without the tryptophan residue (86). Therefore, the tryptophan residue at the entrance of tunnel is thought to have a key function in crystalline cellulose degradation.

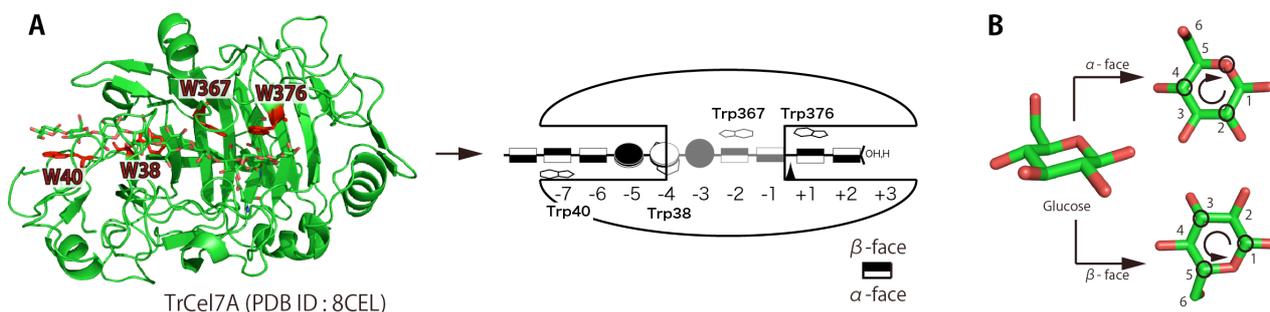


Fig. 1-14 Tryptophan residues along subsites of *TrCel7A* (A) and glucose surface (B)

	37	38	40	49	103	179	181	367	376																																																																																								
<i>Trichoderma reesei</i> Cel7A	QS	A	G	K	N	V	G	A	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	Y	A	N	M	L	W	L	D	S	T																																																					
<i>Acremonium thermophilum</i> Cel7A	QA	A	C	L	T	A	E	T	H	P	S	L	Q	W	Q	K	T	A	P	G	--	S	C	T	T	V	S	G	V	T	I	D	A	N	W	R	L	H	T	A	N	S	S	T	--	N	C	A	-	G	K	N	V	G	A	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	Y	A	N	M	L	W	L	D	S	T						
<i>Acremonium thermophilum</i> Cel7B	Q	G	V	G	T	Q	T	E	T	H	P	K	L	T	F	O	K	S	A	A	G	--	S	C	T	T	T	Q	N	G	E	V	I	D	A	N	W	R	L	H	D	K	N	G	V	T	--	N	C	S	Y	A	T	N	I	G	S	R	M	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	Y	S	M	L	W	L	D	S	T				
<i>Aspergillus aculeatus</i> CBH	Q	G	V	G	T	Q	T	E	T	H	P	S	L	T	W	Q	T	S	S	G	--	S	C	T	T	S	S	V	I	D	A	N	W	R	L	H	T	V	H	E	V	G	G	V	T	--	N	C	S	-	G	K	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	L	S	I	W	D	D	H	A	D	M	L	W	L	D	S	T						
<i>Aspergillus nidulans</i> CBH	Q	K	V	G	T	Q	G	A	E	V	H	P	L	T	W	Q	T	S	S	G	--	S	C	T	T	V	N	G	E	V	I	D	A	N	W	R	L	H	T	V	H	G	G	V	T	--	N	C	S	-	G	K	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	L	S	I	W	D	D	H	A	D	M	L	W	L	D	S	T						
<i>Aspergillus niger</i> CBH-A	Q	A	G	L	T	E	E	V	H	P	L	T	W	Q	K	T	S	E	G	--	S	C	T	E	O	S	G	S	V	I	D	A	N	W	R	L	H	S	V	N	D	S	T	--	N	C	S	-	G	K	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	L	S	I	W	D	D	Y	A	S	M	E	W	L	D	S	T								
<i>Aspergillus niger</i> CBH-B	Q	G	V	G	T	Q	T	E	T	H	P	S	L	T	W	Q	T	S	D	G	--	S	C	T	T	N	D	G	E	V	I	D	A	N	W	R	L	H	T	S	S	A	T	--	N	C	S	-	S	K	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	L	S	I	W	D	D	Y	A	A	D	M	L	W	L	D	S	T							
<i>Chaetomium thermophilum</i> CBH-1	Q	A	C	S	L	T	T	E	T	H	P	R	L	T	W	K	R	C	T	S	G	--	N	C	S	T	V	N	G	A	V	T	I	D	A	N	W	R	L	H	T	V	S	G	S	T	--	N	C	S	-	G	H	T	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	N	M	L	W	L	D	S	T			
<i>Chrysosporium lucknowense</i> Cel7A	Q	N	A	C	L	T	A	E	N	H	P	L	T	W	S	K	C	T	S	G	--	S	C	T	S	V	Q	G	S	I	T	I	D	A	N	W	R	L	H	T	R	S	A	T	--	N	C	S	-	G	Y	S	T	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	V	N	M	L	W	L	D	S	T			
<i>Cochliobolus carbonum</i> Cel1	Q	G	V	G	T	A	E	N	H	P	L	T	W	Q	T	C	T	G	T	G	T	G	G	T	N	C	S	N	S	G	S	V	I	D	A	N	W	R	L	H	A	H	N	V	G	G	V	T	--	N	C	S	-	F	S	S	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	V	N	M	L	W	L	D	S	T
<i>Fusicoccum</i> sp. CBH	Q	G	V	G	T	Q	T	A	E	T	H	P	K	L	T	Q	K	C	T	T	A	G	--	G	C	T	D	Q	S	T	I	V	L	D	A	N	W	R	L	H	T	V	D	G	Y	T	--	N	C	S	-	G	T	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	A	N	L	W	L	D	S	T				
<i>Humicola grisea</i> Cel7A	Q	G	V	C	S	L	T	E	T	H	P	L	T	W	S	K	C	T	S	G	--	C	S	D	V	K	G	S	V	I	D	A	N	W	R	L	H	T	Q	T	S	G	S	T	--	N	C	S	-	P	Y	S	K	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	S	N	M	L	W	L	D	S	T				
<i>Humicola grisea</i> CBH-2	Q	A	C	S	L	T	T	E	R	H	P	L	T	W	S	K	C	T	A	G	--	C	Q	C	T	V	Q	A	S	I	T	L	D	S	N	W	R	L	H	T	Q	S	G	S	T	--	N	C	S	-	G	Y	S	T	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	S	N	M	L	W	L	D	S	T		
<i>Hypocrea koningii</i> CBH	Q	A	G	T	I	A	E	N	H	P	R	M	T	W	K	R	C	S	G	P	G	--	N	C	Q	T	V	Q	G	E	V	I	D	A	N	W	R	L	H	-	N	N	G	--	N	C	S	-	E	Y	G	T	N	I	G	S	R	L	Y	L	Q	C	A	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	S	N	M	L	W	L	D	S	T					
<i>Irpex lacteus</i> Cel1	Q	S	A	C	L	T	Q	S	E	T	H	P	L	T	W	Q	K	C	S	S	G	--	T	C	T	Q	O	T	G	S	V	I	D	A	N	W	R	L	H	T	A	N	S	S	T	--	N	C	S	-	A	-	G	K	N	V	G	A	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	Y	A	N	M	L	W	L	D	S	T			
<i>Irpex lacteus</i> Cel2	Q	O	V	G	T	Q	M	A	E	V	H	P	L	T	W	S	K	C	T	S	G	--	C	T	N	O	N	T	A	V	L	D	A	N	W	R	L	H	T	T	S	G	V	T	--	N	C	S	-	G	T	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	S	M	O	W	L	D	S	T						
<i>Melanocarpus albomyces</i> Cel7B	Q	O	V	G	T	N	O	A	E	N	H	P	L	T	W	S	K	C	T	A	S	G	--	C	T	T	S	S	V	L	D	A	N	W	R	L	H	T	H	T	T	G	Y	T	--	N	C	S	-	G	T	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	L	S	I	W	D	D	H	A	N	M	L	W	L	D	S	T							
<i>Penicillium chrysogenum</i> Cel7A	Q	R	A	G	N	E	T	P	E	N	H	P	L	T	W	Q	R	C	T	A	P	G	--	N	C	O	T	N	A	E	V	I	D	A	N	W	R	L	H	D	-	D	N	M	O	--	N	C	S	-	E	Y	G	T	N	I	G	S	R	L	Y	L	Q	C	A	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	N	M	L	W	L	D	S	T			
<i>Penicillium funiculosum</i> Cel7A	Q	G	V	G	T	S	T	A	E	V	H	P	L	T	W	Q	K	C	T	A	G	--	S	C	T	S	O	S	G	V	I	D	A	N	W	R	L	H	S	V	N	D	S	T	--	N	C	S	-	G	K	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	H	A	N	M	L	W	L	D	S	T						
<i>Penicillium janthinellum</i> Cel7A	Q	O	I	G	T	A	E	T	H	P	S	L	S	W	S	T	C	K	S	G	--	S	C	T	T	N	S	G	A	I	L	D	A	N	W	R	L	H	V	G	N	V	T	S	A	T	--	N	C	S	-	G	S	N	I	G	S	R	L	Y	L	Q	C	P	R	D	L	K	F	I	N	G	M	V	L	V	M	S	L	I	W	D	D	Y	S	V	N	M	L	W	L	D	S	T			
<i>Phanerochaete chrysosporium</i> Cel7C	Q	G	V	G	T	N	A	E	N	H	R	L	T	S	Q	K	C	T	S	G	--	G																																																																											

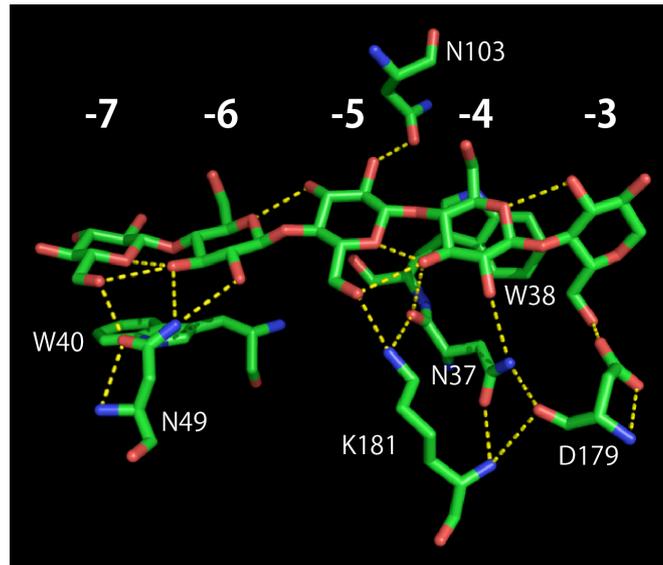


Fig. 1-16 Interactions between cellulose chain and unique residues for GH7 CBH in *TrCel7A*

1.3.3 Comparison of loop structures between CBH from ascomycete and basidiomycete

Phylogenetic tree of the GH7 enzymes was made from the alignment of sequences classified as 'characterized' enzymes in the CAZy database (Fig. 1-17). As a result, Cel7s have been clearly divided into two major sub-families (CBH and EG), and also CBHs were divided into the two groups; one comprising the ascomycete CBHs and the other comprising the basidiomycete CBHs. In Fig. 1-17, the differences of the loop regions among the three enzyme groups are also shown. Ascomycete CBHs have four loop regions, but EGs have only two of them and their subsites are more open. On the other hand, basidiomycete CBHs lack one loop region covering the active site compared with the other CBHs. Thus, the structure of basidiomycete CBHs are in between of the structures of ascomycete CBHs and EGs.

Another difference between basidiomycete and ascomycete about GH7 enzymes is their secretion pattern. Ascomycete typically produces both CBH and EG type GH7 enzymes in cellulolytic culture, whereas basidiomycete produces only similar CBH type GH7 enzymes. For example, white-rot fungi *Phanerochaete chrysosporium* produces

*PcCel7C* and *PcCel7D* to a cellulolytic culture. Structural comparison of *PcCel7C* and *PcCel7D* showed no difference in the amino acids that are expected to interact directly with the cellulose chain in the subsites of the catalytic domains (76). However, three amino acids constructing loop regions, located in loop2 and loop4 covering the active-site tunnel, are different, as shown in Fig. 1-18. This information implies that the possibility to clarify the importance of loop regions for processivity by the activity comparison of CBHs from ascomycete and basidiomycete.

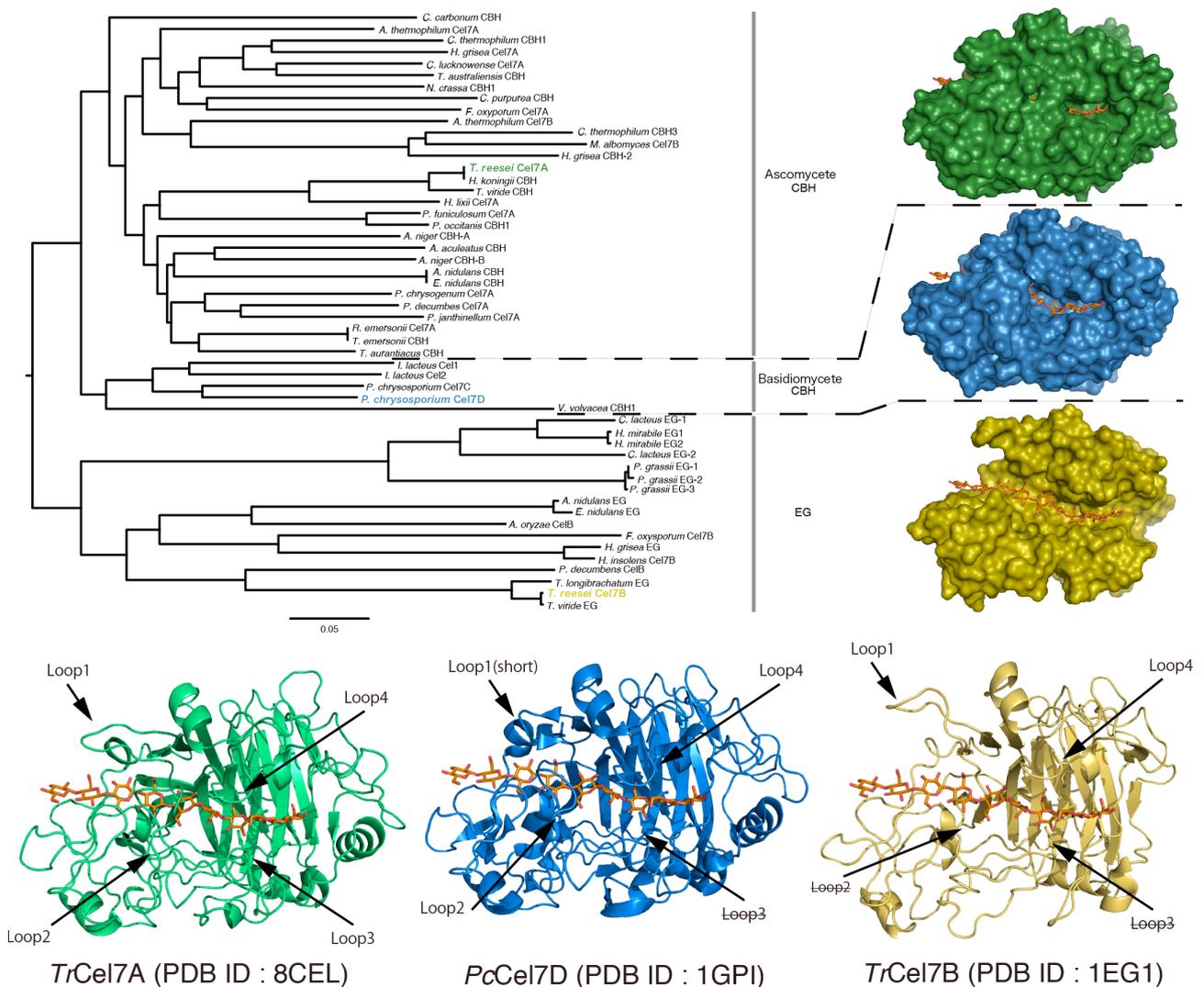


Figure 1-17 Phylogenetic tree of the sequence based alignment of Cel7s classified as 'characterized' at the CAZy database and differences of loop regions among three groups.

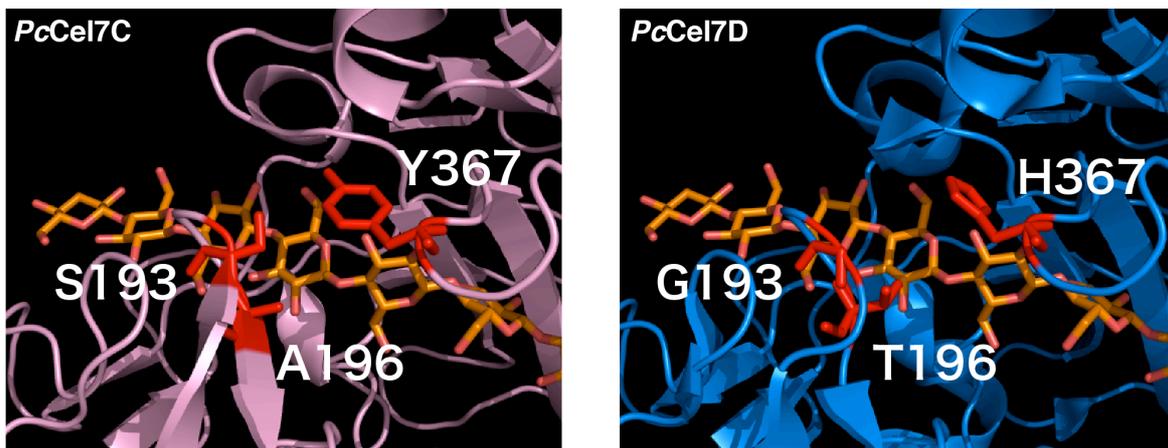


Fig. 1-18 Different amino acids around the loop regions between *PcCel7C* and *PcCel7D*.

Y367 and H367 residues are located on the loop4 and the others are located on the loop2.

#### 1.3.4 Recent advances in GH7 cellobiohydrolase research

Because cellulose is an insoluble substrate, GH7 CBH should act on the liquid-solid interface. In case of GH7 CBH with CBD, the reaction is started by the adsorption on the cellulose surface. After binding on the surface by CBD, catalytic core catches the cellulose chain end and takes it into the catalytic center. Finally, GH7 CBH start the processive reaction and desorb from the cellulose (Fig. 1-19). All the steps are important for understanding the CBH reaction. Therefore, research history about the each step of reaction (Adsorption, substrate intake and processive reaction) and kinetics of GH7 CBH are summarized below.

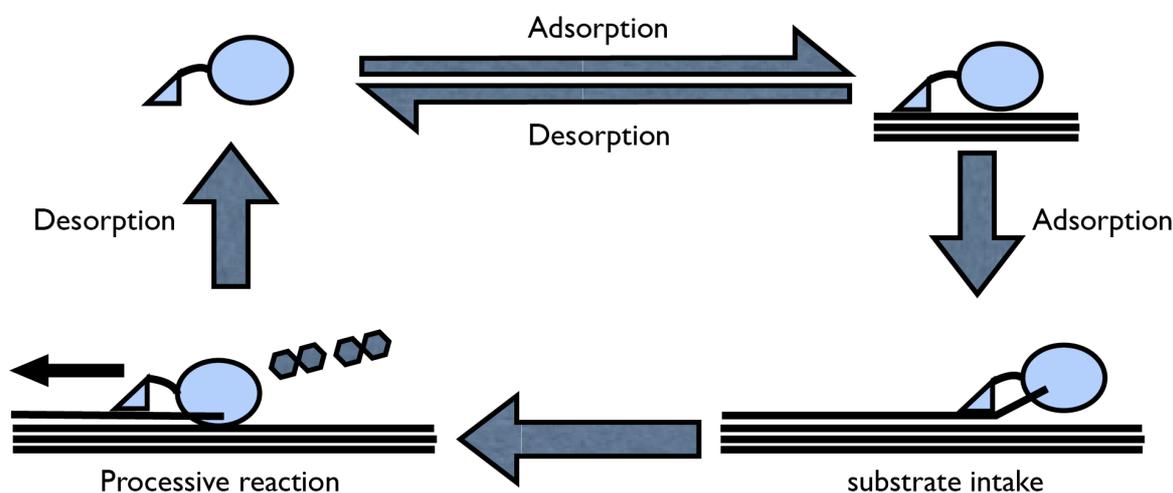


Fig. 1-19 Reaction cycle of GH7 cellobiohydrolase

## **Adsorption**

Adsorption of purified CBH I on the cellulose was reported in 1977, and visualized by electron microscopy in 1983 (87,88). The shape of CBH I (head and long tail) and the two-domain structure was revealed by small angle X-ray scattering in 1986 (89). To investigate the effect of the small domain, the activities of intact CBH I and catalytic core of CBH I (papain treated CBH I) were compared in 1986 and 1988 (90,91). As a result, it was shown that C-terminal domain of CBH I is related to the adsorption and degradation of crystalline cellulose. In addition, three dimensional structure of C-terminal domain of CBH I has been solved by NMR and adsorption potential was suggested by the structure in 1989 (81). Substrate recognition by CBD has been reported in detail in 1997 (82). Furthermore, the reversibility of adsorption of *TrCel7A* and the half time of adsorption of CBD from *TrCel7A* was reported in 1999 (92,93). The double binding model of CBD and CD has been suggested in 1991 due to the two domain structure and the curve fitting of the adsorption data to the Langmuir's two-binding-site model (94). Although the two binding model has been used for a long time, the new model of binding: negative cooperative adsorption was proposed in 2012 (95). The model was proposed because the binding isotherm of a fusion protein of the CBD of *TrCel7A* with red-fluorescent protein (RFP) even fitted better to the Hill's model than to the Langmuir's single site model. In 2013, one of the role of CBD was studied at high substrate concentration by comparing the activity of enzymes with and without CBD under (96). It was suggested that CBMs are crucial for substrate recognition only at low substrate consistencies.

## **Substrate intake**

GH7 CBH was first defined as an exo-glucanase, because it was not capable of decreasing the viscosity of CMC, and its specificity towards the reducing end of cellulose was reported in 1989 (45,46). However the possibility of endo attack of GH7 CBH to

cellulose was also reported in 1993 (97). In 1995, although the disadvantage of GH7 CBH for endo attack was shown by a comparison of the structures of GH7 CBH and EG, the possibility of endo attack remained. Additionally, the endo activity of GH7 CBH to both end labeled cellotetraose was suggested in 1997 (98). Recently, the probability of endo attack of *TrCel7A* to reducing-end labeled amorphous cellulose or bacterial cellulose was reported in 2011 (99).

### **Processive reaction**

After the X-ray structure of catalytic core of CBH I (*TrCel7A*) was solved in 1994 and the structures of *TrCel7A* and EG I (*HlCel7*) were compared in 1995, the possibility of processive reaction of CBH was suggested due to the unique tunnel-like structure of CBH (49,51). The processive reaction and the specificity to reducing end of *TrCel7A* (CBH I from *Trichoderma viride*) was first visualized by an electron microscopy in 1998 (100). Also in 1998, the definition of processivity (processivity = the velocity of moving / the velocity of dissociation) was suggested and, it was concluded that the processivity was limited by the non-productive binding enzymes and the roughness of cellulose surface (101). Additionally, the processivity of *TrCel7A* to avicel was calculated by the production ratio of cellobiose per glucose, and the value was estimated to be between 5 and 10 in 1998 (102). In 2003, the relationship between the loop structure and processivity was shown by the comparison of *TrCel7A* wild type (WT), loop region deleted mutant of *TrCel7A* and GH7 CBH from a basidiomycete *Phanerochaete chrysosporium* (*PcCel7D*) (103). The rough indexes of processivity were calculated from the production ratio between cellobiose and the sum of glucose and cellotriose. In the paper, the processivity of *TrCel7A* WT on bacterial microcrystalline cellulose (BMCC; bacterial cellulose (BC) treated by hydrochloric acid) was 23. In 2005, the processivity of *TrCel7A* was estimated by the separation of products from initial attack and processive reaction with reducing end labeled BC and

BMCC (104). The estimated processivities were about  $42 \pm 10$  and  $88 \pm 10$  on labeled BC and BMCC respectively. In the same year, the subsite affinities of *TrCel7A* catalytic domain were calculated (105). As a result, it was shown that the subsites energy gradient was suitable for substrate sliding into the catalytic center, and the ability of processive reaction of *TrCel7A* was supported. Recently, the processive movements of *TrCel7A* molecules were directly observed by high-speed atomic force microscopy (HS-AFM) in 2009 (73). In 2011, the processivity of *TrCel7A* on reducing end reduced BC was reported to be  $61 \pm 14$ , and the theoretical processivity, which was calculated from catalytic constant ( $k_{cat}$ ) and dissociation rate constant ( $k_{off}$ ) under the hypothesis;  $k_{cat} \gg k_{off}$  and all enzymes were making enzyme-substrate complex, was reported to be  $4000 \pm 570$  (99). In 2012, the observed processivity of *TrCel7A* to  $^{14}\text{C}$ -labeled BC was reported to be  $66 \pm 7$  (106). In the study by Jalak et al. (106), the reducing-end of BC was not labeled or modified, but it was assumed that the initiations of processive reaction of each *TrCel7A* molecule were limited to one by the adding of amorphous cellulose into the reaction mixture to trap the free enzymes.

### **Kinetics and reaction model**

For a long time, the importance of (i) the character of the enzyme, (ii) the physical properties of the substrate and (iii) the interactions between enzyme and substrate for understanding of cellulase reaction has been discussed (Fig. 1-20). About the first factor, many studies have been reported as described above. The first kinetics of GH7 CBH to soluble cello-oligosaccharides was reported as Michaelis-Menten equation with product inhibition in 1980 (107). About the factors of cellulose, the surface area and the accessibility of cellulase, the degree of crystallinity and degree of swelling were suggested around 1970 and reviewed in 1980 (108). In 1975, the first kinetic model of crude cellulase, which includes the enzyme adsorption as the Langmuir's single adsorption

model, was proposed (109).

The first kinetic study about purified GH7 CBH on cellulose was reported in 1977 (87). In the paper, kinetic model of GH7 CBH was also described as a combination of Michaelis-Menten equation and the Langmuir's single adsorption, but the substrate concentration was described by the surface area of  $\alpha$ -cellulose ( $m^2 / g$ -cellulose). Additionally, the competitive inhibition of cellobiose and inhibition constant were reported. In 1987, the kinetics of GH7 CBH and GH6 CBH and six endo-glucanase was reported (110). In the paper, the surface area of avicel was described by the amount of maximum protein adsorption ( $mg$ -protein /  $mg$ -cellulose), and the relationship between the hydrolysis velocity and the amount of adsorbed enzyme was discussed. In 1998, the decrease of velocity at the early stage of cellulose degradation was discussed by the data of the cellobiose production velocity vs amounts of adsorbed protein and free protein (101). As a result, it was suggested that the non-productive binding enzyme and the crumbled surface of cellulose hinder cellulase activity. The amount of produced cellobiose vs reaction time was fitted to a dual exponential equation;  $p(t) = a \cdot (1 - e^{-b \cdot t}) + c \cdot (1 - e^{-d \cdot t})$ , but the meaning of the equation was not discussed. In 2006, the concept of surface density was introduced to cellulase kinetics (111). In this concept, cellulose surface area was defined by maximum amount of adsorbable enzyme calculated from Langmuir's two-binding-site model. The

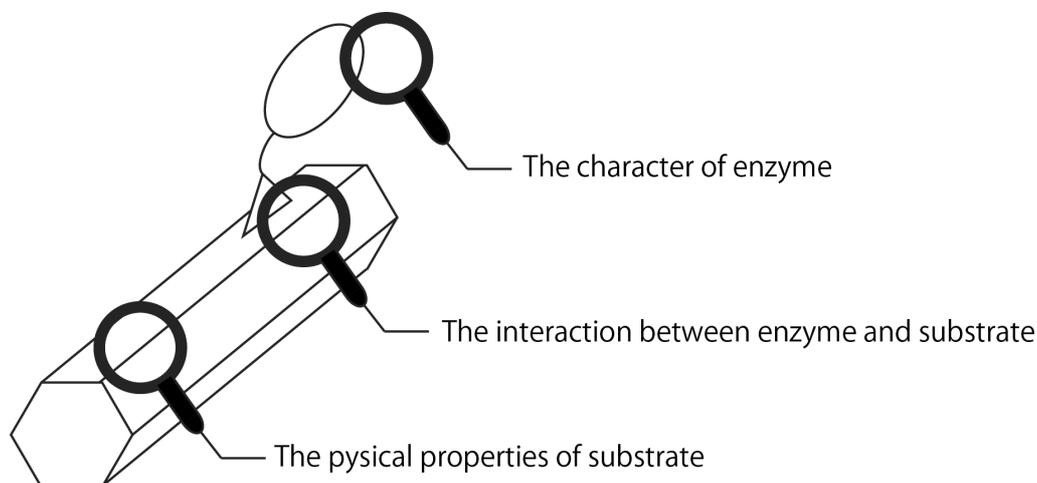


Fig. 1-20 The factors for analysis of cellulase reaction

density was calculated from the amount of adsorbed enzyme per the maximum potential value of surface amount of adsorbed enzyme. Based on the velocity per adsorbed enzyme and surface density, it was suggested that the congestion of enzyme molecules on cellulose surface decrease the activity of each *TrCel7A* molecule. In addition, the reason of high activity of *TrCel7A* to cellulose III was clarified by surface density analysis in 2007 (112). As a result, it was shown that cellulose III has larger hydrophobic surface and higher accessibility for cellulase. The hypothesis of 'the traffic jam' of *TrCel7A* has been supported by observation via HS-AFM in 2011 (113). In 2010, the amount of enzyme, which containing the cellulose chain in subsite, was calculated from the degree of inhibition on soluble substrate hydrolysis by cellulose (114). In the paper, the importance of  $k_{off}$  was suggested under the hypothesis;  $k_{cat} \gg k_{off} \cdot n$  ( $n$  is processivity). In a paper of 2012, the mean of double exponential equation was estimated by the measurement of subsites occupied enzyme on  $^{14}\text{C}$ -labeled BC (106). By the result, it was concluded that the division of equation is caused by the small amount of enzymes, which shows different observed processivity or different velocity.

## 1.4 Aim of this thesis

Although the X-ray structure of GH7 CBH was solved already in 1994, there has not been enough discussion was achieved about the relationship between activity and enzyme structure. Therefore, the aim of this thesis is to investigate the relationship between the ability of CBHs to degrade crystalline cellulose efficiently and the unique structure of the catalytic domain of GH7 CBH (Fig. 1-21).

The tryptophan residues and the tunnel-like structure are the unique features of the structures of CBH. The tryptophan residues seems to be related to the initial substrate recognition and to the intake of the cellulose chain, because the residues are located at the entrance of the active site tunnel. To analyze the roles of these tryptophan residues in the substrate recognition mechanism, the activities for soluble model substrate, cello-oligosaccharides and crystalline cellulose were compared among the wild type (WT) and tryptophan mutants (W38A and W40A) of *TrCel7A*. In addition, the processive reactions of CBH molecules from *T. reesei* (*TrCel7A*) and *P. chrysosporium* (*PcCel7C* and *PcCel7D*) on the crystalline cellulose were observed using the high-speed atomic force microscopy. The relationship between processive reaction and tunnel like structure constructed by loop regions were analyzed by comparing of the kinetic parameters of the CBHs.

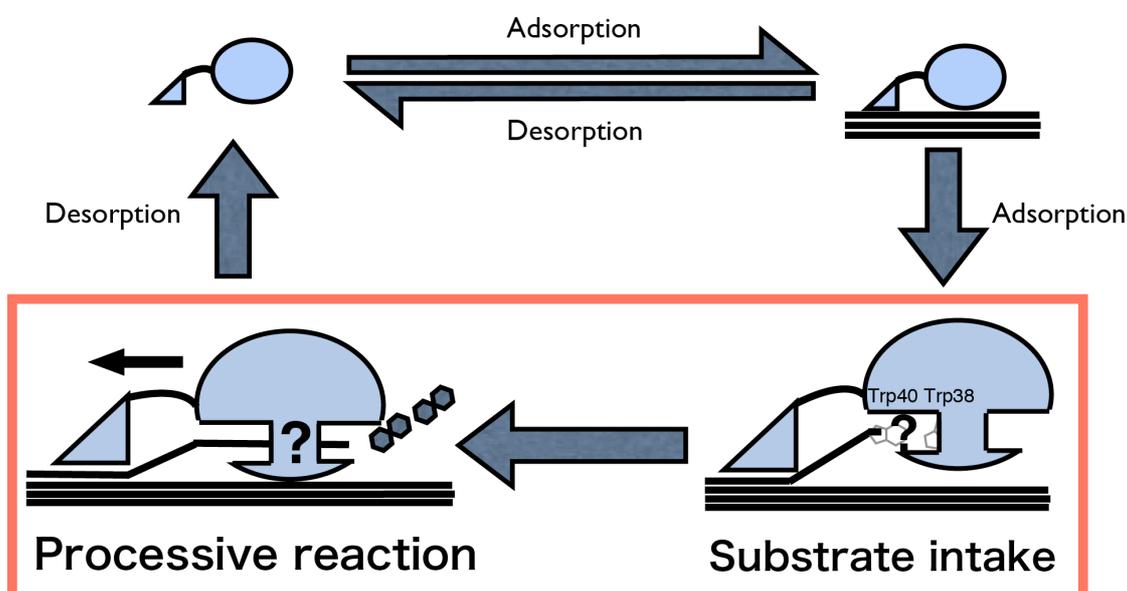


Fig. 1-21 Research target of this thesis

## **2. Materials and methods**

## 2.1 Enzyme production

### 2.1.1 *TrCel7A* mutants

*Trichoderma reesei* strains, in which each of mutated *TrCel7A* genes were introduced, were cultivated in a Chemap LF 20 fermenter (working volume 15 l). The cultivation medium contained 60 g of whey powder, 30 g of spent grain, 5.0 g of  $\text{KH}_2\text{PO}_4$  and 5 g of  $(\text{NH}_4)_2\text{SO}_4$  per liter of medium. Inocula (spore suspensions in 50% glycerol stored at  $-80^\circ\text{C}$ ) were grown on the same medium with only  $20\text{ g l}^{-1}$  carbon source and  $10\text{ g l}^{-1}$  spent grain, buffered with  $15\text{ g l}^{-1}$   $\text{KH}_2\text{PO}_4$ , in two stages of  $4 \times 50\text{ ml}$  (72 hours, 200 rpm,  $28^\circ\text{C}$ ) and  $5 \times 200\text{ ml}$  (10% v/v transfer, 48 hours, 200 rpm,  $28^\circ\text{C}$ ). The fermentation lasted for 192 hours and the other cultivation conditions were:  $T = 28^\circ\text{C}$ , pH 4–5 (lower limit controlled by addition of  $\text{NH}_4\text{OH}$ , upper limit by addition of  $\text{H}_3\text{PO}_4$ ), dissolved oxygen (DO)  $> 30\%$  (agitation 400 – 800 rpm), aeration  $5\text{ l min}^{-1}$ . Foaming was controlled by automatic addition of Struktol J633 polyoleate antifoaming agent (Schill & Seilacher, Germany). After fermentation, the mycelium was separated by centrifugation for 30 min at 4000 g,  $4^\circ\text{C}$ .

### 2.1.2 *PcCel7C* and *PcCel7D*

*Phenerochaete chrysosporium* strain K-3 was grown on 2 L of Kremer and Wood medium containing 2% cellulose (CF11; Whatman) as the sole carbon source in a jar fermenter with 5 L working volume (Takasugi Seisakusho, Tokyo, Japan). The temperature was  $37^\circ\text{C}$  and the pH was maintained at 5.0 with phosphoric acid and potassium hydroxide. Stirring was done at 300 rpm and air was supplied at  $2.0\text{ l/min}$ . After cultivation for 120 hours, the culture supernatant was separated using a glass filter membrane. Extracellular proteins were precipitated with ammonium sulfate at 70% saturation and stored at  $4^\circ\text{C}$ .

## 2.2 Enzyme purification

### 2.2.1 *TrCel7A* and *TrCel7B* and *TrCel7A* mutants

*TrCel7A* WT was purified from Celluclast® 1.5L (Novozymes, Denmark). The crude enzyme was desalted by a 50 ml column volume (C.V.) of gel filtration column (TOYOPEARL® HW-40s: TOSOH, Japan), which was equilibrated with 20 mM potassium phosphate buffer pH 7.0. The protein concentrations of each fraction was measured by Protein Assay (Bio-rad, USA), and the fractions containing protein were mixed and injected to an anion exchange column (TOYOPEARL® DEAE-650s, C.V. = 150 ml), which was equilibrated with the same buffer. The proteins were eluted by a linear gradient of 300 mM potassium chloride. After the measurement of protein concentration, the hydrolysis activity for *p*-nitrophenyl- $\beta$ -D-lactoside (*p*NPL: Sigma-Aldrich, USA) was measured, and purity of each fraction was estimated by SDS-PAGE. The fractions containing about 60 kDa enzyme, which has activity for *p*NPL, were collected and mixed with half amount of 2.0 M ammonium sulfate solution. The enzyme solution was injected to a hydrophobic interaction column (TOYOPEARL® Phenyl-650s, C.V. = 70 ml), which was equilibrated with 20 mM of sodium acetate buffer pH 5.0 containing 1.0 M ammonium sulfate. The enzyme was eluted by a linear reverse-gradient of ammonium sulfate. The protein concentration, activity for *p*NPL and purity of each fraction was analysed by the same methods, and the enzyme was collected. The buffer solution of the enzyme was changed to 20 mM of sodium acetate buffer pH 5.0 by ultra-filtration. The enzyme solution was filtrated by 0.1  $\mu$ m of a PVDF membrane (Ultrafree Centrifugal Filters: Merch-Millpore, Deutschland). Enzyme concentration was calculated from the absorbance at 280 nm. Molar absorbance coefficient of *TrCel7A* WT was 83370 M<sup>-1</sup>·cm<sup>-1</sup>. Two tryptophan mutants were also purified from each enzyme solution by the same methods, but the value of molar absorbance coefficient was 77680 M<sup>-1</sup>·cm<sup>-1</sup>. Purified enzymes were stored at 4 °C. *TrCel7B* was also purified from Celluclast® 1.5L. The flow-through fraction of the DEAE-650s was collected,

and the buffer was changed to 20 mM Tris-HCl buffer pH 8.0. The enzyme solution was injected to a DEAE-650s (C.V. = 150 ml) column equilibrated with the same buffer, and enzymes were eluted by a linear gradient of 300 mM sodium chloride. The fractions contained 60 kDa enzyme, which has activity for *p*NPL, were mixed with half amount of 2.0 M ammonium sulfate solution. The enzyme solution was injected to a Phenyl-650s column (C.V. = 70 ml) equilibrated with 20 mM sodium acetate buffer pH 5.0 containing 1 M ammonium sulfate. The enzyme was collected, and the buffer was changed to 20 mM Tris-HCl buffer pH 8.0. The enzyme solution was injected to a DEAE-650s (C.V. = 150 ml) column equilibrated with the same buffer, and enzymes were eluted by a linear gradient of 150 mM sodium chloride. The buffer of the TrCel7B was changed to the same buffer as the other enzymes, and the enzyme concentration was calculated by the same method. Used value of the molar absorbance coefficient was 70590 M<sup>-1</sup>•cm<sup>-1</sup>.

### 2.2.2 *PcCel7C* and *PcCel7D*

Precipitate of extracellular proteins was dissolved in 20 mM potassium phosphate buffer (pH 7.0) and was applied to a DEAE-Toyopearl 650s column (C.V.: 150 ml) equilibrated with 20 mM potassium phosphate buffer (pH 7.0) after desalting. Proteins were eluted with a linear gradient 0 to 0.3 M KCl in 1050 ml. The protein concentration of fractions were estimated by Protein Assay kit according to the manufacture's instruction. Twenty  $\mu$ l of samples were incubated with 20  $\mu$ l of 10 mM *p*-nitrophenyl lactoside (*p*NPL), 20  $\mu$ l of 1M sodium acetate buffer (pH 5.0) and 140  $\mu$ l of water. Reactions were stopped by the addition of 20  $\mu$ l of 2 M Na<sub>2</sub>CO<sub>3</sub> and measured absorbance of the reaction mixture at 405 nm ( $\epsilon_{405} = 16608 \text{ M}^{-1}\text{cm}^{-1}$ ). Their *p*NPL hydrolyze activity units were calculated using the standard curve of *p*-nitrophenol. The purity of each fraction was checked by SDS-PAGE and collected fractions containing about 58-kDa and 62-kDa proteins with *p*NPL-hydrolyzing activity. The 62-kDa protein in flow through fractions were equilibrated against

20 mM Tris-HCl buffer, pH 8.0, and applied to a SuperQ-Toyopearl 650S (TOSOH; C.V.: 100 ml) equilibrated with same buffer. Proteins were eluted with a linear gradient 0 to 0.1 M NaCl in 800 ml of 20 mM Tris-HCl, pH 8.0. After that, ammonium sulfate was added to protein solution to final concentration 1 M and fractionated on a Phenyl-Toyoperl 650s column (C.V.: 67 ml) equilibrated with 20 mM sodium acetate buffer (pH 5.0) containing 1 M ammonium sulfate. Proteins were eluted with a reverse linear gradient 1 to 0 M ammonium sulfate in 134 ml. Finally, protein solution was equilibrated against 20 mM Tris-HCl buffer (pH 8.0) and fractionated on a DEAE-Toyopearl 650S column (C.V.: 150 ml) equilibrated with same buffer. Proteins were eluted with a linear gradient 0 to 0.05 M NaCl in 1350 ml. The fractions were assayed as described above and fractions containing 62-kDa enzyme with *p*NPL-hydrolyzing activity were collected and equilibrated with 20 mM sodium acetate buffer (pH 5.0). The 58-kDa protein solution was fractionated on a Phenyl-Toyoperl 650s column (TOSOH; C.V.: 67 ml) using the same method. Finally, protein solution was equilibrated against 20 mM potassium phosphate buffer (pH 7.0). Proteins were eluted with a linear gradient 0 to 0.05 M KCl in 1350 ml. The fractions were assayed as described above and fractions containing 58-kDa enzyme and *p*NPL-hydrolyzing activity but not having *p*-nitrophenyl- $\beta$ -D-glucoside-hydrolyzing activity were collected and equilibrated with 20 mM sodium acetate buffer, pH 5.0. The purity was checked by SDS-PAGE analysis on a 12 % polyacrylamide gel.

## 2.3 Cellulose preparation

### 2.3.1 Crystalline cellulose

Highly crystalline cellulose I $\alpha$ -rich (HCC I $\alpha$ ) and III $_1$  were prepared from the green algae *Cladophora* sp. as described previously (24,111). The cells preserved in 1 M sodium hydroxide were washed by water and incubated 2h in the sodium acetate buffer pH 5.0 containing 0.3% sodium chlorite 80°C. Samples were washed by water and incubated in

1.0 M sodium hydroxide solution at 25°C over night. After one more treatment by sodium chlorite, the stalk ends of fascicles were removed and the bundles were cut into moderate length. The sample was incubated in 0.1 M hydrochloric acid at 110°C for 20 min. Treated sample was washed by water and broken into small fragments using a double-cylinder type homogenizer. Cellulose fragments were recovered by centrifugation at 8000 g for 20 min and separated to two groups. For HCC I<sub>α</sub>, 30 g of sample was treated in 180 ml of 4 N hydrochloric acid at 80°C with agitation (500 rpm) for 8h. For preparation of cellulose III<sub>i</sub>, 25 g of sample was suspended in 125 ml of water and frozen by liquid nitrogen. The sample was freeze-dried for 2 days and dried at 50°C over-night. Dried cellulose was treated by ammonia under the supercritical condition for one hour and remained ammonia was volatilized at room temperature. Treated cellulose was also hydrolyzed by 4 N hydrochloric acid under the same condition. Acid hydrolyzed samples were concentrated by centrifugation at 6000 g and re-suspended in water. After ten times washes, pH of the sample was checked and the suspension was centrifuged at 3000 g and removed precipitate. Purified cellulose was stored at 4°C.

### 2.3.2 Amorphous cellulose

Phosphoric acid swollen cellulose (PASC) was prepared from Avicel (Funacell II, Funakoshi, Tokyo, Japan). Avicel was dissolved in 85% (w/w) phosphoric acid and made completely clear by smashing with a grass stick. After over-night incubation at 4°C, cellulose was regenerated in water and homogenized by a high-speed blender. Cellulose suspension was washed by water and stored at 4°C.

## 2.4 Enzyme characterization

### 2.4.1. Activity comparison to cellulose

Purified *TrCel7A* WT and mutants were incubated with purified HCC I<sub>α</sub> in 50 mM sodium acetate buffer pH 5.0 at 30°C for 2 h. Enzyme and cellulose concentration were 0.5 μM and 0.1% respectively. Soluble products were recovered from the reaction mixture by a 0.22 μm PVDF filter (MultiScreen HTS filter plate: Merck-Millipore, Deutschland) and separated by a HPLC (Jasco, Japan) equipped with a NH2P-4E column (Shodex, Japan) with a linear gradient by water and acetonitrile. The amount of glucose to cellohexaose was quantified by a Corona CAD detector (Thermo Fisher Scientific, USA) with standard curves of each oligosaccharide.

*TrCel7A*, *PcCel7C* and *PcCel7D* (2.0 μM) were incubated with 0.1 % of cellulose I<sub>α</sub>, cellulose III<sub>1</sub> or PASC in 50 mM sodium acetate buffer pH 5.0 at 30°C. After 15, 30, 60 and 120 min (and additionally 5, 10 min for cellulose I<sub>α</sub>) incubation, reaction mixture was filtrated by a MultiScreen 0.22 μm filter plates and products were analyzed by the same methods. Produced sugar (Glucose to cellotriose) concentrations in the reaction mixture were fitted to the equation,  $q(t) = a \cdot (1 - e^{-bt}) + c \cdot t$ , and a differential equation,  $v(t) = a \cdot b \cdot e^{-bt} + c$  (from cellulose III<sub>1</sub>) and  $q(t) = a \cdot (1 - e^{-bt}) + c \cdot (1 - e^{-dt})$ , and a differential equation,  $v(t) = a \cdot b \cdot e^{-bt} + c \cdot d \cdot e^{-dt}$  (from cellulose I<sub>α</sub> and PASC), where  $a$ ,  $b$ ,  $c$ , and  $d$  are constants, and  $t$  is the incubation time, in order to calculate the production velocity.

### 2.4.2 Determination of kinetic parameters to *p*NP-cellobioside

Purified *TrCel7A* WT and mutants were incubated with various concentration of *p*-nitrophenyl-β-D-cellobioside (*p*NPC: Sigma-Aldrich, USA) in 50 mM sodium acetate buffer pH 5.0 at 30°C for 120 min. Enzyme concentration was 0.1 μM. The reaction was stopped by addition of 20 μl of 2 M sodium carbonate to 200 μl of mixture. The amount of released *p*-nitrophenol was detected as the absorbance at 405 nm by a V-660 spectrophotometer

(Jasco, Japan) and the concentration was calculated from standard curve of *p*-nitrophenol. In case of *PcCel7C* and D, enzyme concentration was 1.0  $\mu$ M and incubation time was 10 min. The plots of velocity vs *p*NPC concentration were fitted to Michaelis-Menten equation with or without substrate inhibition ( $v = k_{cat} \cdot E_0 \cdot S / (K_m + S + S^2 / K_s)$  or  $v = k_{cat} \cdot E_0 \cdot S / (K_m + S)$ ).

#### 2.4.4 Comparison of kinetic parameters to cellotetraose

Various concentrations of cellotetraose were incubated with the enzymes in 20 mM sodium acetate buffer pH 5.0 at 30°C. The incubation times of WT, W38A were 30 min, W40A was 60 min. The Enzyme concentration to 10  $\mu$ M and 25  $\mu$ M of substrate was 0.0125  $\mu$ M, 50 to 100  $\mu$ M was 0.025  $\mu$ M and equal to or higher than 150  $\mu$ M was 0.05  $\mu$ M. The methods of separation and quantification of products were same as above. All plots were fitted by the global fitting program packaged in Igor Pro 6 (WaveMetrics, Portland, USA) under the restrictions calculated from the reaction models.

#### 2.4.5 Molecular dynamics simulation of catalytic domains of WT and W40A

The initial coordinates of the *TrCel7A* CD were taken from Protein Data Bank (PDB) structure 8CEL (84), 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 ([http://nbc222.ucsd.edu/pdb2pqr\\_1.8/](http://nbc222.ucsd.edu/pdb2pqr_1.8/)), 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 WTcat and W40Acat, 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<sup>+</sup> counter ions were added to obtain electrostatic neutrality. All the simulations were

performed using the Amber 11 package (115). The Amber ff03 force field was used for proteins. We employed the general Amber force field with the AM1-BCC partial charges for the N-terminal pyroglutamic acid, 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 (116) and the long-range electrostatic forces were calculated using the particle-mesh Ewald method (117). The trajectory analysis was conducted by using the Amber module ptraj, and the snapshot structures were visualized with VMD (118). 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 CD of WT and W40A.

#### 2.4.6 High-speed AFM observation

Movements of molecules of the three cellobiohydrolases were observed by high-speed atomic force microscopy based on previous reports (113,119). 0.1% suspension of cellulose III<sub>I</sub> was dropped on a grass stage covered by highly oriented pyrolytic graphite (HOPG) and incubated for 10 min. After rinsing the unbounded cellulose with 50 mM sodium acetate buffer pH 5.0, the stage was set on the instrument, and 70  $\mu$ l of the same buffer was added. 2  $\mu$ l of 20  $\mu$ M cellulase solution was added and moving molecules, which start and end movement in the observation area were analyzed. The averaged moving velocities of each molecule were drawn by IGOR Pro 6 and fitted to Gaussian distribution. Adsorption times of molecules were calculated from the number of observed frames, and the plots of number of molecules to each range of adsorption time were fitted to the equation:  $f(t) = a \cdot e^{-bt}$ , which  $a$  and  $b$  is constants,  $t$  is adsorption time. Highly crystalline cellulose III<sub>I</sub> was used as substrate.



## **3. Results and discussion**

### 3.1 Analysis of cellulose recognition by CBH unique tryptophan residues

#### 3.1.1 Comparison of kinetic parameters to *p*NP-cellobioside

To quantify the differences among the WT and mutated enzymes, the activities to soluble model substrate were compared. The velocities of *p*NP release from each *p*NPC concentration were plotted in Fig. 3-1. Although W40A showed substrate inhibition in the tested range of *p*NPC concentration, the plots of other enzymes fitted to the normal Michaelis-Menten equation. The kinetic parameters of each enzyme are summarized in Table 3-1 and estimated reaction model is shown as Scheme 3-1. The  $k_{cat}$  of WT was  $0.418 \text{ min}^{-1}$ , and the value of W40A was slightly lower, but the value of W38A was about 1.4 times higher. The  $K_m$  of WT was  $62.5 \mu\text{M}$  and the value of W40A was a little smaller. The  $K_m$  of W38A was increased up to 14 times higher than that of WT. Therefore, the value of  $k_{cat}/K_m$  of W40A was best, and the value of WT was second. The values of W38A was more than ten times lower than the value of WT. The substrate inhibition constant ( $K_s$ ) of W40A was  $6500 \mu\text{M}$ .

These differences between WT and mutants should be caused by the mutagenesis of tryptophan residues. Additionally, Trp40 and Trp38 residues are located far from the catalytic site, and thus they are not directly affect to the hydrolysis reaction. Therefore, the differences of kinetic parameters imply the changes in substrate-intaking characters, which include the first binding step and the substrate sliding step. The presence of tryptophan residues along the subsite would be better for the substrate binding, but worse for the sliding due to the stabilization of complex. In case of W38A mutant, the value of  $K_m$  was drastically increased and  $k_{cat}/K_m$  was decreased. These results imply that interaction by the Trp38 residue is efficient for the *p*NPC hydrolysis. But the increase of  $k_{cat}$  in the W38A mutant also implies that the strong interaction between Trp38 and substrate becomes less important when the substrate concentration is high enough to intake *p*NPC without Trp38 residue. On the other hand, W40A showed higher value of  $k_{cat}/K_m$  than WT. This change

was caused by the increase of the value of rate constant  $k_1$ . In addition, a big difference between W40A and the other enzymes is substrate inhibition. This difference can't be explained by a simple substrate inhibition model, because the change of  $K_s$  needs the changes of plus side of subsites in that model. One possibility, which can explain the results, is the productive-nonproductive complex model. In case of *p*NPC hydrolysis, the faces of substrate should be considered because the two different enzyme-substrate complex (productive and non-productive) can be formed (Scheme 3-1). In this model,  $K_s$  is affected by the rate constants about the substrate intake.

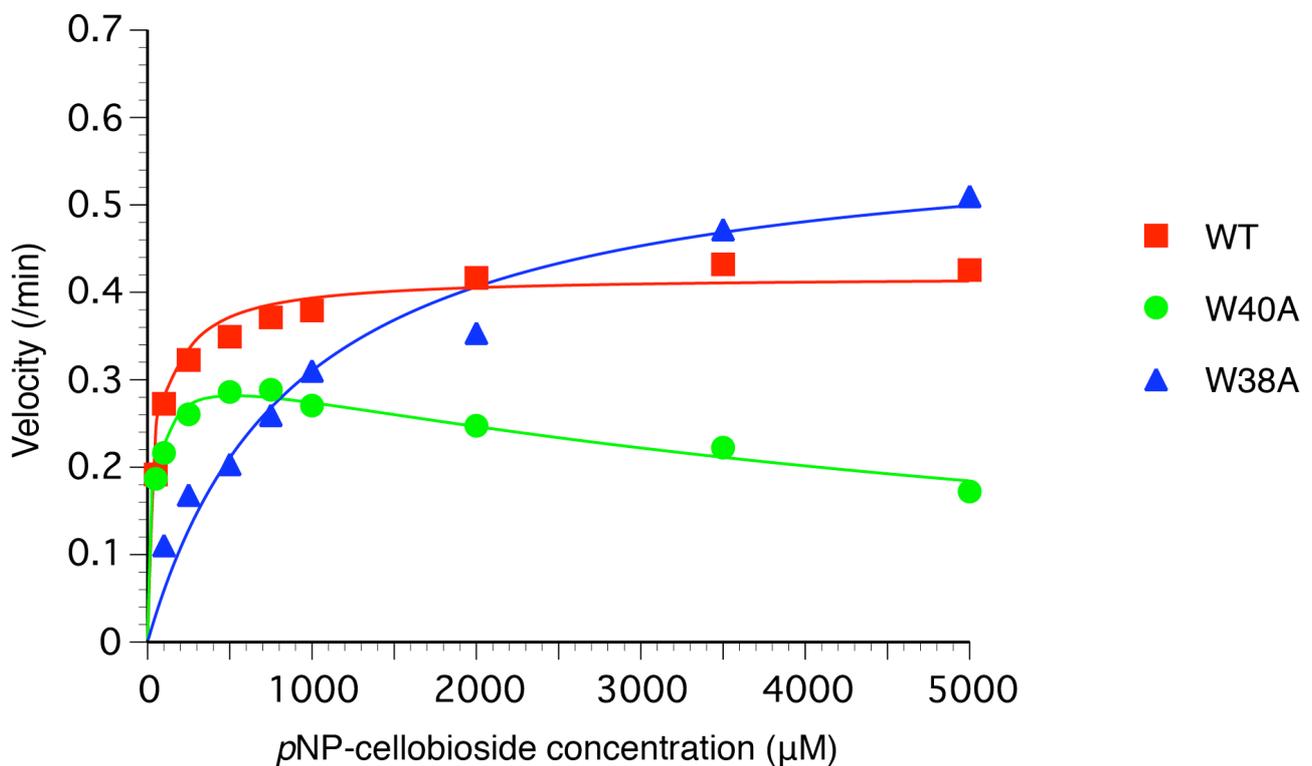


Fig. 3-1 *p*NP-cellobioside hydrolysis by *Tr*Cel7A WT and tryptophan mutants

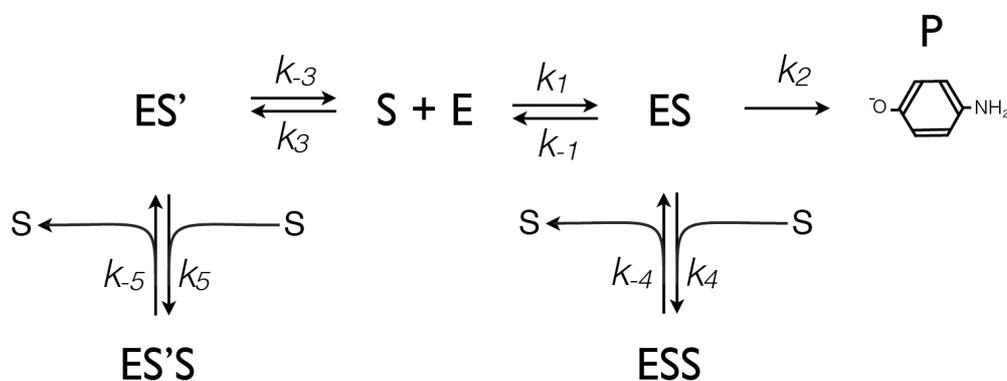
Enzyme concentration: WT, W40A and W38A was 0.1 μM

Reaction conditions: 120 min incubation in 50 mM NaAc pH5.0 at 30°C

Table 3-1 Kinetic parameters of *Tr*Cel7A WT and mutans to *p*NP-cellobioside

Enzyme	$k_{cat}$ (min <sup>-1</sup> )	$K_m$ (μM)	$k_{cat}/K_m$ (min <sup>-1</sup> μM <sup>-1</sup> )	* $K_s$ (μM)
WT	0.418 ± 0.008	62.5 ± 8.2	66.9 × 10 <sup>-4</sup>	-
W40A	0.328 ± 0.012	43.0 ± 6.9	76.3 × 10 <sup>-4</sup>	6500 ± 1020
W38A	0.589 ± 0.033	898 ± 164	6.6 × 10 <sup>-4</sup>	-

\* $K_s$  is a substrate inhibition constant



$$\frac{1}{E_0} \cdot \frac{d[pNP]}{dt} = \frac{k_{cat} \cdot [S]}{K_m + [S]} = \frac{\frac{k_1 k_2 k_{-3}}{k_1 k_{-3} + (k_{-1} + k_2) k_3} \cdot [S]}{\frac{(k_{-1} + k_2) k_{-3}}{k_1 k_{-3} + (k_{-1} + k_2) k_3} + [S]} \quad (\text{WT and W38A})$$

$$\frac{1}{E_0} \cdot \frac{d[pNP]}{dt} = \frac{k_{cat} \cdot [S]}{K_m + [S] + \frac{[S]^2}{K_s}} = \frac{\frac{k_1 k_2 k_{-3}}{k_1 k_{-3} + (k_{-1} + k_2) k_3} \cdot [S]}{\frac{(k_{-1} + k_2) k_{-3}}{k_1 k_{-3} + (k_{-1} + k_2) k_3} + [S] + \frac{[S]^2}{\frac{k_{-4} k_{-5} \{k_1 k_{-3} + (k_{-1} + k_2) k_3\}}{k_1 k_{-3} k_4 k_{-5} + (k_{-1} + k_2) k_3 k_{-4} k_5}}} \quad (\text{W40A})$$

$$\frac{k_{cat}}{K_m} = \frac{k_1 k_2}{k_{-1} + k_2}$$

Scheme 3-1 Reaction model of *p*NPC hydrolysis of *Tr*Cel7A

### 3.1.2 Comparison of kinetic parameters to cellotetraose

To clarify the differences among the three enzymes, kinetic parameters for cellotetraose were determined. At first, production velocities of three enzymes from various concentration of cellotetraose are shown in Fig. 3-3. WT and W40A produced only glucose, cellobiose and cellotriose, but W38A additionally produced cellohexaose from the high concentration of cellotetraose. WT and W40A showed apparent decreases of velocity under high concentration of cellotetraose, but W38A showed only slight decrease of production velocities of glucose and cellotriose. These plots of velocity vs cellotetraose concentration were fitted to the equations, which were calculated from the reaction model of *TrCel7A* (Scheme 3-2). This model includes two types of cellotetraose intake from minus side of subsite. In case of the substrates longer than cellotriose, both types cause productive complex (hydrolysis points were different). Additionally, the model includes trans-glycosylation and the intake from the plus side of subsite and/or endo-intake.

The determined parameters are summarized in Table 3-2. Especially for cellotetraose, the products patterns depended on the initial intake pattern. Therefore, the values of  $k_{cat1}/K_m$  or  $k_{cat4}/K_m$  reflect the efficiency of each substrate intake and hydrolysis. The  $k_{cat1}/K_m$  of W40A ( $0.69 \mu\text{M}^{-1}\text{min}^{-1}$ ) was 1.1 times higher than WT ( $0.61 \mu\text{M}^{-1}\text{min}^{-1}$ ) and 1.8 times higher than W38A ( $0.38 \mu\text{M}^{-1}\text{min}^{-1}$ ). Additionally, the  $k_{cat4}/K_m$  of W40A ( $0.48 \mu\text{M}^{-1}\text{min}^{-1}$ )

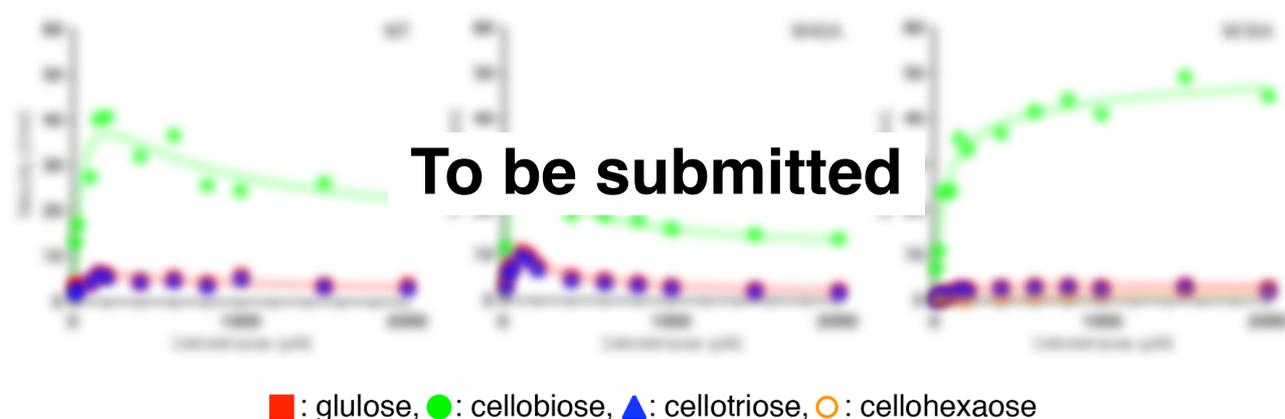
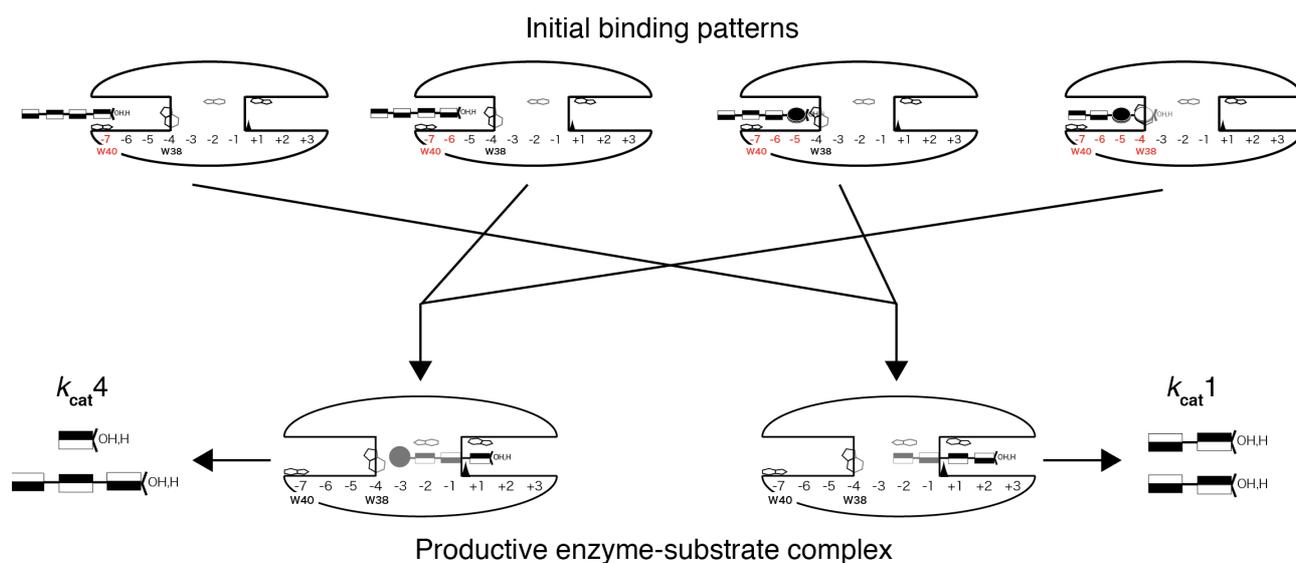


Fig. 3-3 cellotetraose hydrolysis by *TrCel7A* WT and mutants



was 2.3 times higher than WT ( $0.21 \mu\text{M}^{-1}\text{min}^{-1}$ ) and 8.2 times higher than W38A ( $0.06 \mu\text{M}^{-1}\text{min}^{-1}$ ). These results show that Trp38 residue increases the efficiency, but Trp40 residue decreases the efficiency of cellotetraose hydrolysis. The ratio of these two efficiency ( $k_{\text{cat}1}/k_{\text{cat}4}$ ) of WT was 3.0, and the values of W40A and W38A were 1.4 and 6.5 respectively. Thus, these two tryptophan residues have different tendency for the substrate-intaking pattern. This tendency can be explained by the relationship between initial binding patterns and productive complex (Scheme 3-3). Trp40 residue relates to all of the initial binding patterns, and Trp38 residue relates only to one pattern. When Trp40 residue is mutated to alanine, all patterns will be affected, but the effect of decrease of affinity of the initial binding for production of glucose and cellotriose can be covered by Trp38 residue. In case of W38A mutant, the affinity in the initial binding pattern, which forms the productive complex for glucose and cellotriose, may be decreased.

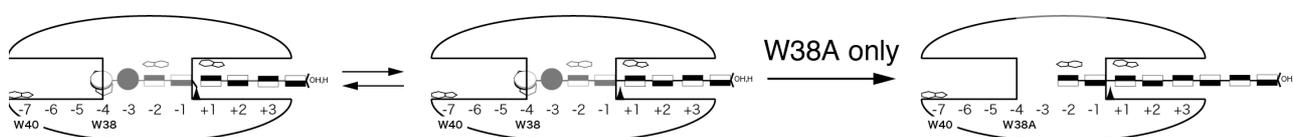
Another difference among the three enzymes was production of cellohexaose. The production of cellohexaose indicates the trans-glycosylation activity of W38A mutant, but



Scheme 3-3 Relationship between initial binding patterns and productive complex

structure of W38A around the catalytic site are same as WT and W40A. One hypothesis, which can explain the difference, is forming the cellooctaose and the hydrolysis of it (Scheme 3-4). In this model, production of cellohexaose is also depend on the affinity at subsite -4. Although WT and W40A hydrolyse almost all cellooctaose by the strong interaction of Trp38 residue, W38A mutant can't keep the end of cellooctaose at subsite -4, and result in the production of cellobiose and cellohexaose. This model is supported by the decrease of processivity shown in the crystalline cellulose degradation and cello-oligosaccharides hydrolysis.

The last matter is the possibilities of substrate intake from plus side subsite and endo-intake. These possibilities are necessary for the curve fitting of the three enzymes, because observed cellobiose production velocities of them are higher than the expected values from the simple substrate inhibition model without square terms of substrate on the numerators. Unfortunately, these possibilities can't be separated in the equation, but contributions of them to the cellobiose production velocities are also larger in the higher concentration of substrate. The rate constant of these reactions ( $k_{cat2}$ ) of W40A and WT were higher than that of W38A, and they showed negative correlation with  $K_{m2}$ . Thus, the reaction is regulated by the binding of second substrate. The productive binding of second substrate always competes with the binding of first substrate around the catalytic site. This relationship means that the substrate intake from both minus and plus side are inhibited by each other. Therefore, in case of WT and W40A, small  $K_{m2}$  values suggested that cellotetraose intaken from the minus subsite were trapped on the Trp38 residue, and additional substrate can easily bind to the catalytic center.



Scheme 3-4 Trans-glycosylation model of *TrCel7A* W38A mutant

### 3.1.3 Activity comparison to crystalline cellulose

To clarify the importance of Trp40 and Trp38 residues, the activities of the mutants of the two residues to crystalline cellulose were compared with the activity of WT and endoglucanase (*TrCel7B*). The quantified concentrations of products are shown in Fig. 3-5. Only glucose and cellobiose were detected in the reaction solution of all enzymes. *TrCel7A* WT produced about 5.8  $\mu\text{M}$  of glucose and 60  $\mu\text{M}$  of cellobiose from HCC I $_{\alpha}$ , but two mutants produced about 10  $\mu\text{M}$  of cellobiose. On the other hand, *TrCel7B* produced about 5.5  $\mu\text{M}$  of glucose and 27  $\mu\text{M}$  of cellobiose at the same condition. In addition, the ratio of cellobiose and glucose (= rough index of processivity) of WT was 10.3, but that of W40A and W38A were 6.5 and 7.3 respectively. These results shows that the interactions by both two tryptophane residues are important for degradation of crystalline cellulose and processive reaction of CBH. Therefore, analysis of the reaction for the cellulose molecule chain on the crystalline cellulose surface is needed to clarify the role of Trp40 residue.

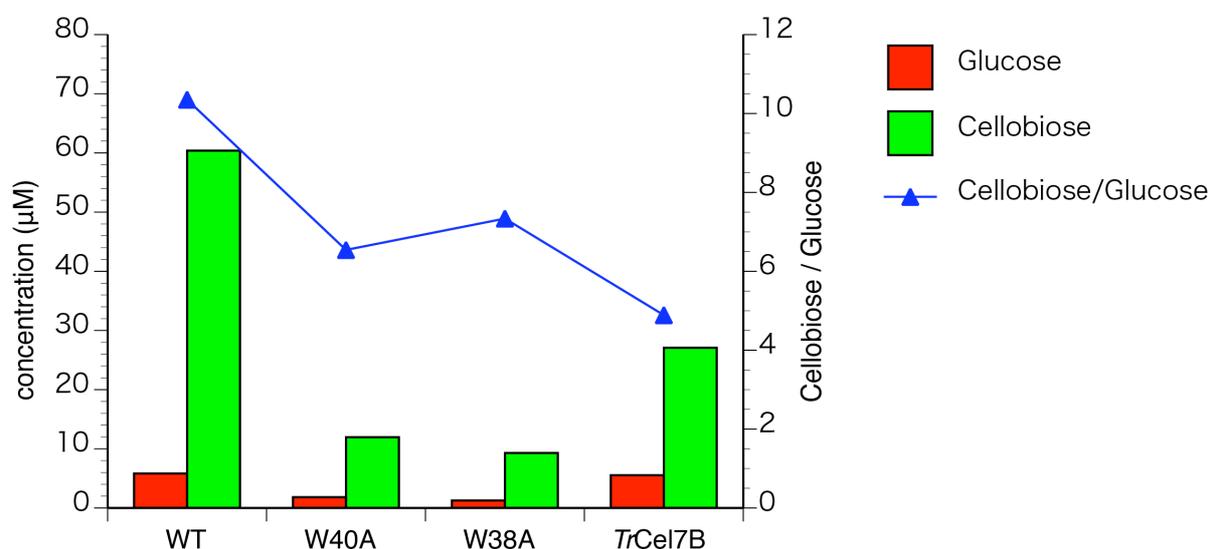


Fig. 3-5 Comparison of crystalline cellulose degradation activity

Red is produced glucose and green is cellobiose.

Substrate: 0.1% of high crystalline cellulose I $_{\alpha}$  from *Cladophora sp.*

Reaction condition: 0.5  $\mu\text{M}$  enzyme was incubated with substrate for 2h in 50 mM NaAc pH 5.0 at 30°C

### 3.1.4 Estimation of the role of Trp40 residue by molecular dynamics simulation

Quantification of activity for insoluble molecular chain of cellulose on the crystalline cellulose surface is difficult, because the quantification of true-active enzyme in the reaction mixture is impossible by using normal biochemical experiments. Difference between crystalline cellulose and soluble oligosaccharides are the length of chains and interactions between chains. The interactions between soluble oligo-saccharides are weak, whereas the molecular chains of cellulose on an insoluble crystal are strongly packed. This difference suggests that the flexibility of the chains on the crystalline surface is lower than that of soluble oligo-saccharides. Accordingly, the role of Trp40 residue in catching a less accessible chain ends can be evaluated. The effect of accessibility of the cellulose chain ends to the substrate intake were estimated by the molecular dynamics simulation.

Results of the simulation of WT and W40A are shown in Fig. 3-6 and Fig. 3-7. By these simulations, the difference between WT and W40A was clarified. WT could keep and intake the substrate from both subsites -7 and -5, but W40A released the chain end from

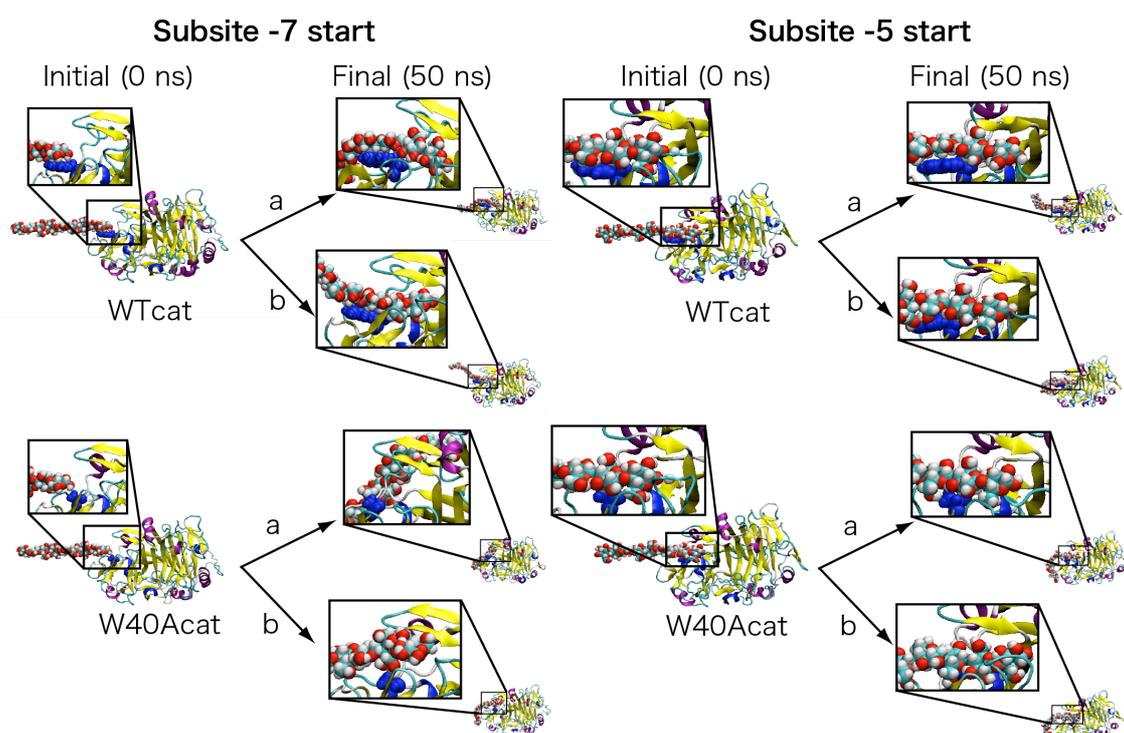


Fig. 3-6 Results of molecular dynamics simulation of CD of WT and W40A

subsite -7. These results imply that at least three glucosidic residues should be free from the crystalline cellulose surface for the substrate intake without Trp40 residue. The possibility of dissociation of three glucosidic residues is the cubic of dissociation possibility of one glucosidic residue. Therefore, it is reasonable that the activity of W40A to crystalline cellulose was smaller than WT.

As a result, it was revealed that both Trp40 and Trp38 residues are related to the substrate intake of *TrCel7A*. Trp40 residue catches the reducing-end glucose ring of cellulose chain on the crystal, and Trp38 keeps the chain and sends it into catalytic center processively (Scheme 3-6). These two residues, which are located at the entrance of the tunnel, are specially important for crystalline cellulose degradation, and it is reasonable that they are only conserved among CBHs. In addition, it was suggested that both residues increase the processivity of *TrCel7A*. This effect is also reported as one of the factors, which increases the efficiency of crystalline cellulose degradation. But it is not clearly known why the processive reaction is important for crystalline cellulose degradation.

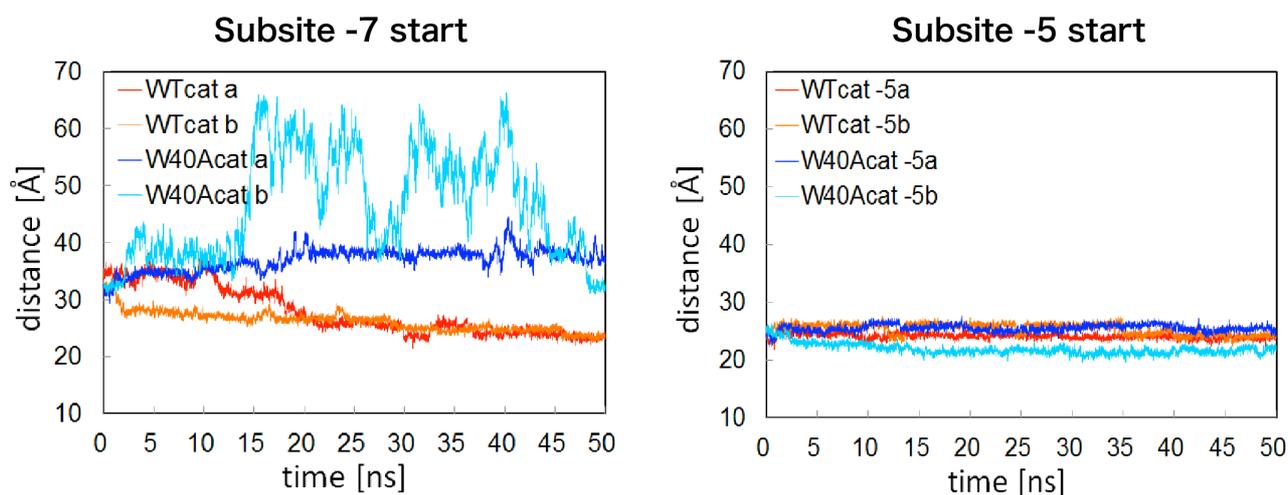
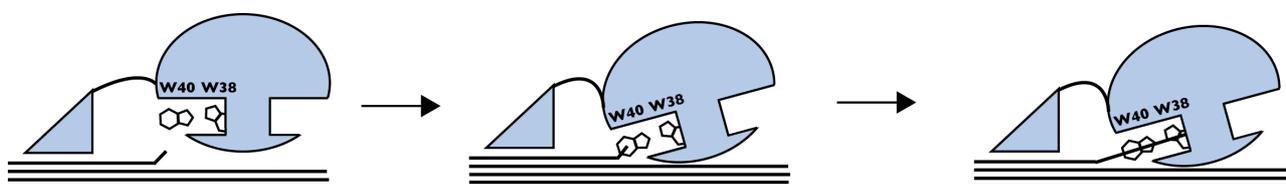


Fig. 3-7 Time course of distances between reducing chain end and catalytic residue



Scheme 3-6 Estimated roles of CBH unique tryptophan residues

## 3.2 Analysis of processive reaction of GH7 cellobiohydrolases

### 3.2.1 General characterization of GH7 CBHs from ascomycete and basidiomycete

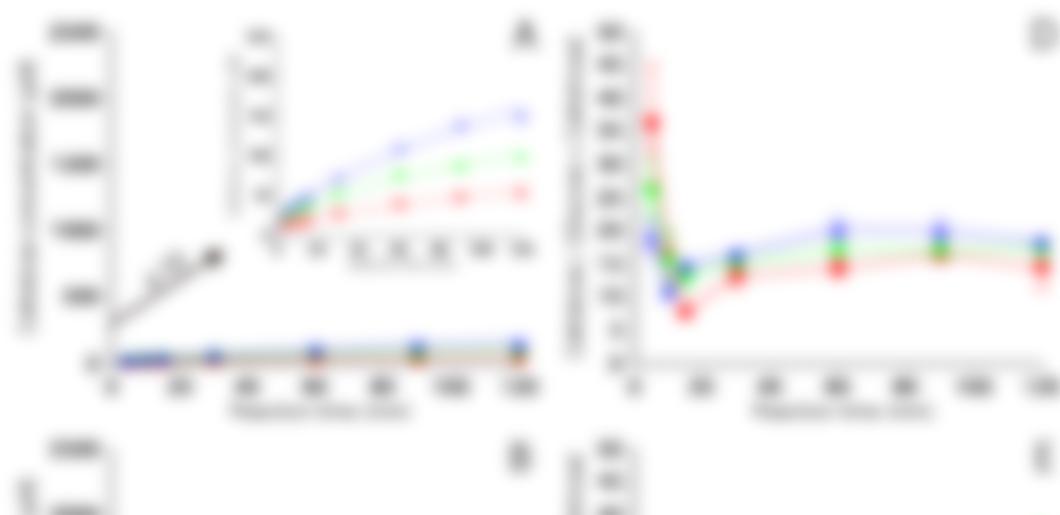
The importance of processivity for crystalline cellulose degradation has been discussed since the x-ray structures of CBH and EG were solved. But it is still unclear why processivity is important for the crystalline cellulose degradation, because the determination of processivity of cellulase is difficult. Recently the processive movement of a single molecule of *TrCel7A* was directly observed by high-speed atomic force microscopy (73). Therefore, comparison of the processive movements of CBHs from an anamorph of ascomycete *Trichoderma reesei* (*TrCel7A*) and basidiomycete *Phanerochaete chrysosporium* (*PcCel7C* and *PcCel7D*) may clarify the relationship between the structure of loop regions and processivity. In addition, the combination of single molecule analysis and biochemical measurement of activity for cellulosic substrate may show the importance of processivity for crystalline cellulose degradation.

Before the HS-AFM observation, the general characters of GH7 CBHs from *TrCel7A*, *PcCel7C* and *PcCel7D* were biochemically analyzed. Cellobiose inhibition constants to *pNPC* hydrolysis were summarized in Table 3-4. The  $K_s$  value of *PcCel7C* was the highest, and that of *TrCel7A* was the lowest. Then, purified enzymes were also incubated with crystalline cellulose I<sub>α</sub>, III<sub>I</sub> and amorphous cellulose (phosphoric acid swollen cellulose; PASC), and the amounts of products were measured by a HPLC. The progress curves were compared as shown in Fig. 3-8A to C. All enzymes mainly produced cellobiose, but glucose was also detected in the products from all of the substrates. Additionally cellotriose was produced from PASC. *TrCel7A* showed the highest activity for crystalline cellulose I<sub>α</sub> (1.5  $\mu\text{M}/\text{min}$  at 15 min incubation and 0.7  $\mu\text{M}/\text{min}$  at 120 min), and the second was *PcCel7D* (1.0  $\mu\text{M}/\text{min}$  at 15 min incubation and 0.4  $\mu\text{M}/\text{min}$  at 120 min). The activity of *PcCel7C* was apparently the worst among them (0.5  $\mu\text{M}/\text{min}$  at 15 min incubation and 0.2  $\mu\text{M}/\text{min}$  at 120 min). The rate of cellobiose production by *TrCel7A* from crystalline cellulose

III<sub>1</sub> (17.4  $\mu\text{M}/\text{min}$  at 15 min incubation and 10.5  $\mu\text{M}/\text{min}$  at 120 min) was similar to that of *PcCel7D* (16.8  $\mu\text{M}/\text{min}$  at 15 min incubation and 9.7  $\mu\text{M}/\text{min}$  at 120 min) at the all time points. *PcCel7C* was lesser active to crystalline cellulose III<sub>1</sub> (9.8  $\mu\text{M}/\text{min}$  at 15 min and 6.3  $\mu\text{M}/\text{min}$  at 120 min). On the other hand, *PcCel7C* produced cellobiose faster than *PcCel7D* and *TrCel7A* when PASC was used as a substrate. In PASC degradation, the production rate of cellobiose at 15 min incubation by 2.0  $\mu\text{M}$  *PcCel7C* was 34.8  $\mu\text{M}/\text{min}$ , which is almost 1.7 times faster than that of *PcCel7D* (20.4  $\mu\text{M}/\text{min}$ ) and 1.8 times faster than that of *TrCel7A* (19.2  $\mu\text{M}/\text{min}$ ). The ratio between cellobiose and the sum of glucose and cellotriose produced by *TrCel7A* from cellulose I <sub>$\alpha$</sub>  was 20.3, and cellulose III<sub>1</sub> was 34.8 at 60 min (Fig. 3-8 D to F). The ratio of *PcCel7D* from I <sub>$\alpha$</sub>  was 17.2 and from III<sub>1</sub> was 33.3. The values of *PcCel7C* from I <sub>$\alpha$</sub>  was 14.5 and from III<sub>1</sub> was 29.9. In case of PASC degradation, *PcCel7D* showed the highest value (17.5 at 60 min), and *TrCel7A* and *PcCel7C* showed similar values (*TrCel7A*: 12.8 and *PcCel7C*: 13.5).

Table 3-4 kinetic parameters for pNPCellobioside

Enzyme	$k_{cat}$ ( $\text{min}^{-1}$ )	$K_m$ ( $\mu\text{M}$ )	$k_{cat}/K_m$ ( $\text{min}^{-1}\mu\text{M}^{-1}$ )
<i>PcCel7C</i>	$3.42 \pm 0.04$	$491.2 \pm 22.2$	$6.96 \times 10^{-3}$
<i>PcCel7D</i>	$3.82 \pm 0.07$	$965.4 \pm 49.8$	$3.96 \times 10^{-3}$
<i>TrCel7A</i>	$0.418 \pm 0.008$	$62.5 \pm 8.2$	$6.69 \times 10^{-3}$



This figure is same as Figure 2 of the following paper.

Nakamura, A. et al. (2014) Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose. *J. Am. Chem. Soc.*

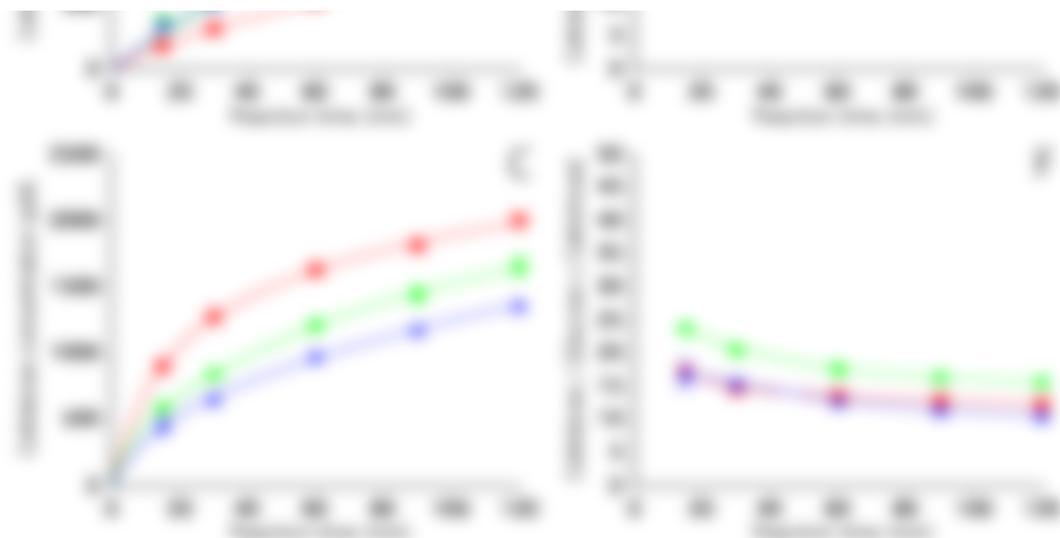


Fig. 3-8 Time course analysis of products amount and ratio of even/odd oligo-saccharides from crystalline cellulose I $_{\alpha}$  (A, D), crystalline cellulose III $_{I}$  (B, E) and PASC (C, F).

0.1% cellulose was incubated with 2.0  $\mu\text{M}$  of enzymes in 50 mM NaAc buffer pH 5.0 at 30°C. Blue triangle: *TrCel7A*, red square: *PcCel7C* and green circle: *PcCel7D*.

### 3.2.2 High-speed AFM observation

*TrCel7A*, *PcCel7C* and *PcCel7D* molecules moving on the surface of crystalline cellulose were observed by high-speed AFM (Fig. 3-9). These results mean that *PcCel7C* and *D* are also processive cellulases as *TrCel7A*. As shown in Fig. 3-9, the moving speeds and distances were different among them. To quantify the differences, the movements of enzyme molecules were analyzed statistically. For the analysis, 176 molecules of *PcCel7C*, 233 molecules of *PcCel7D* and 220 molecules of *TrCel7A* were used. Averaged moving velocities of each molecule were measured and shown in Fig. 3-10. The movement of *TrCel7A* molecules showed a gaussian distribution with center velocity of  $6.8 \pm 3.5$  nm/sec. *PcCel7D* molecules also showed one distribution with the center value:  $9.4 \pm 3.7$  nm/sec. In contrast, *PcCel7C* showed one broad distribution with  $14.7 \pm 9.1$  nm/sec. Examples of moving characters are also showed in Fig. 3-10 with each averaged velocity drawn by line. All enzymes showed stopping states and moving states. Distribution of slopes of *PcCel7C* was the largest among the three enzymes. The longest moving distance



This figure is same as Figure 3 of the following paper.

**Nakamura, A. et al. (2014) Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose. *J. Am. Chem. Soc.***

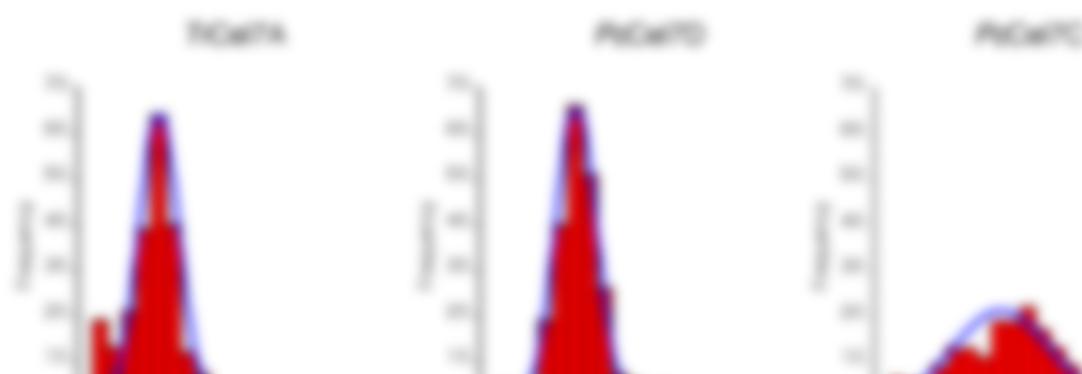


Fig. 3-9 High-speed AFM observation images of three Cel7s.

Positions of each enzyme are marked by arrows.  $0.55 \mu\text{M}$  of enzymes were observed in 50 mM NaAc buffer pH 5.0 at 25°C. Frame rate of the observation was 0.2 frame/sec. Substrate was cellulose III<sub>i</sub>.

of *TrCel7A* was about 250 nm and that of *PcCel7C* and *PcCel7D* was about 120 nm. When the adsorption time of enzyme molecules were analyzed, plots of the three CBHs were all fitted to single exponential decay (Fig. 3-11). The dissociation rate constants ( $k_{\text{off}}$ ) of *TrCel7A* and *PcCel7D* were  $0.20 \pm 0.01 \text{ s}^{-1}$  and  $0.32 \pm 0.04 \text{ s}^{-1}$  respectively. The  $k_{\text{off}}$  of *PcCel7C* ( $0.51 \pm 0.01 \text{ s}^{-1}$ ) was 2.5 times higher than that of *TrCel7A*.

In the AFM observation, only moving molecules with hydrolysis of cellulose chains on the crystalline cellulose surfaces were observed. Therefore, these velocities means how fast processive cycles (Scheme 3-7) were achieved, and the differences in catalytic velocities are directly comparable without the effect of productive/nonproductive ratio of



This figure is same as Figure 4 of the following paper.

**Nakamura, A. et al. (2014) Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose. *J. Am. Chem. Soc.***

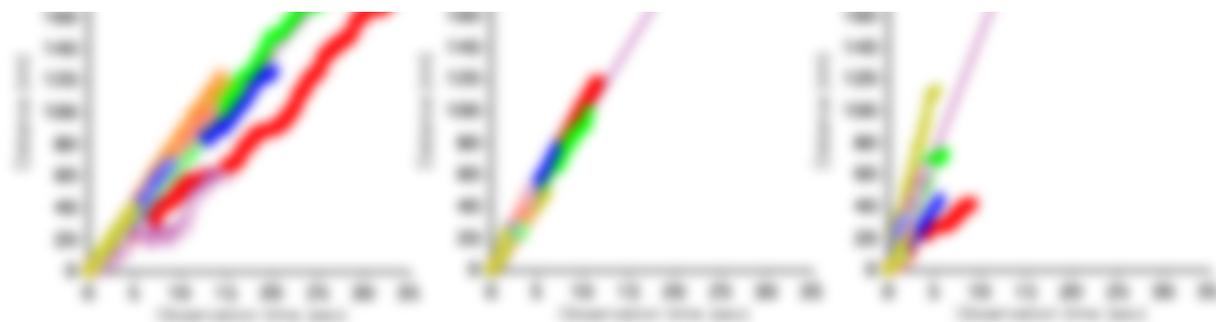


Fig. 3-10 Averaged velocity histograms and examples of moving properties.

Averaged velocities were calculated from the moving-distances and moving-time of each molecule. Peak velocities of each gaussian distribution were drawn in the moving properties by a line.

the enzymes in the reaction mixture. Among the three CBHs, the velocity of *PcCel7C* ( $14.7 \pm 9.1$  nm/sec) was highest and the value of *TrCel7A* was lowest ( $6.8 \pm 3.5$  nm/sec). Because the length of cellobiose is about one nm, these results mean that reaction steps of *TrCel7A* in the processive cycles were slower than *PcCel7C* and D. When the structures were compared among the three enzymes, only *TrCel7A* has loop3 region, which includes the amino acids interact with the cellulose chain at the subsite -2 (Y247), +1 (T246, R251) and +2 (R251) (84). By this difference, it is suggested that *TrCel7A* has stronger affinity to the cellulose chain and the produced cellobiose, and products ejection and chain slide are

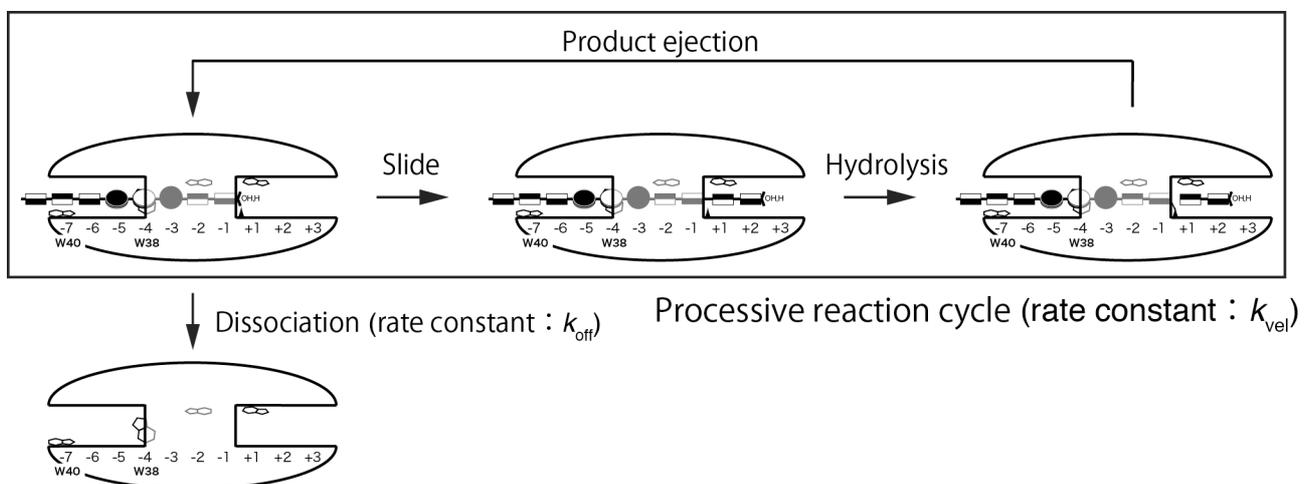
This figure is same as Figure 5 of the following paper.

Nakamura, A. et al. (2014) Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose. *J. Am. Chem. Soc.*



Fig. 3-11 Histograms of adsorption time.

Adsorption times were calculated from how many frames detected during the tracking of the molecules.



Scheme 3-7 Steps in the processive cycle of GH7 CBH

Table 3-5 Summary of AFM observation.

	<i>PcCel7C</i>	<i>PcCel7D</i>	<i>TrCel7A</i>
Dissociation rate constant $k_{\text{off}}$ ( $\text{s}^{-1}$ )	$0.51 \pm 0.01$	$0.32 \pm 0.04$	$0.20 \pm 0.01$
Half-life period (s)	$1.4 \pm 0.0$	$2.2 \pm 0.1$	$3.5 \pm 0.2$
Averaged velocity $k_{\text{vel}}$ ( $\text{nm} \cdot \text{s}^{-1}$ )	$14.7 \pm 9.1$	$9.4 \pm 3.7$	$6.8 \pm 3.5$
Half-life processivity*	19.9	20.3	23.6

\*Half-life processivity = Averaged velocity (nm/s) x Half-life period (s) / cellobiose length (nm)

slower than for the other two enzymes. Especially, It was reported that the catalytic domain of *TrCel7A* has more favorable binding free energy than that of *PcCel7D* (120). Although there is a high overall degree of similarity between *PcCel7C* and *PcCel7D*, three amino acids constructing loop regions of tunnel-like structure are different. They are located in loop2 and loop4 covering the active-site tunnel. Gly193, Thr196, and His367 in *PcCel7D* are replaced by Ser, Ala, and Tyr, respectively, in *PcCel7C*, and it has been predicted that the changes at the loop2 (G193S and T196A) make the loop more flexible (76). These results suggest that higher flexibility of loop2 is related to the faster velocity of processive reaction.

The second difference is the half-life time of adsorption. This value indicates how long the molecules stay on the cellulose surface during the processive reaction. In the AFM observation, only the adsorption by catalytic domain was observed, therefore the distributions of three enzymes were fitted to a single exponential decay. The difference of dissociation rate constants among the three enzymes also means that there is a difference in the affinity to the cellulose chain among the three enzymes. Because *TrCel7A* has more loops than *PcCel7s*, it is reasonable that *TrCel7A* has stronger affinity to cellulose chain. In addition, these results imply the higher affinity of *PcCel7D* than *PcCel7C*, and the flexibility of loop2 is related to the affinity to cellulose chain. Furthermore, half-life processivities of

the three were calculated from averaged velocity and half-life time (Table 3-5). The processivity of *TrCel7A* was the highest, and that of *PcCel7D* was little higher than *PcCel7C*. This order in agreement with the biochemical estimation of processivity, and shows that the length of tunnel-like structures of GH7 CBH reflect the value of processivity as suggested formerly (103). But the values of parameters ( $k_{\text{cat}}$  and  $k_{\text{off}}$ ) related to processivity, are defined by the more detailed structure. The processivity obtained from the ratio of cellobiose/glucose+cellotriose in biochemical experiment reflects the processivity estimated from the HS-AFM observation, suggesting that the biochemical estimation is effective to know rough value of processivity for CBHs. The values of processivity determined by HS-AFM were a little smaller than the values determined by the former biochemical studies (99,104,106). This difference may be caused because HS-AFM needs to apply a force to the molecules for observation, and the substrates are different. Thus, further studies are needed to determine the 'true' processivity.

The processive movement of CBHs from basidiomycete revealed that the regions of loop2 and loop4 are important, but loop3 is not necessary for processive reaction of CBH on crystalline cellulose. Comparing the three CBHs by the HS-AFM observation, *PcCel7D* moves slower than *PcCel7C*, but adsorbs longer than *PcCel7C*. Further, *TrCel7A* moves much slower than *PcCel7D*, but adsorbs much longer than *PcCel7D*. These results imply that moving rate and adsorption time have a trade-off relationship. In other words, hydrolysis rate and dissociation rate constants are in a proportional relation. Both product ejection and chain sliding were estimated to be the rate limiting steps of processive reaction, but the effect of chain slide in the catalytic tunnel should be larger, because sliding needs the re-arrangements of enzyme-cellulose chain interaction along the all subsites, and the energy barrier of it is larger than the other.

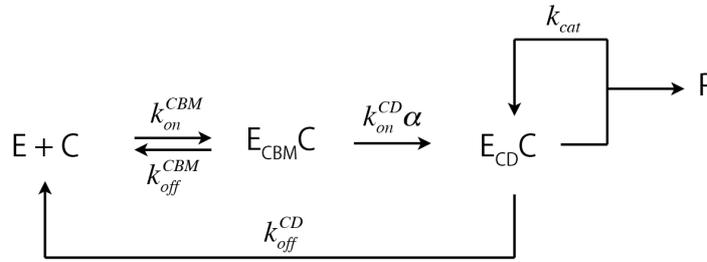
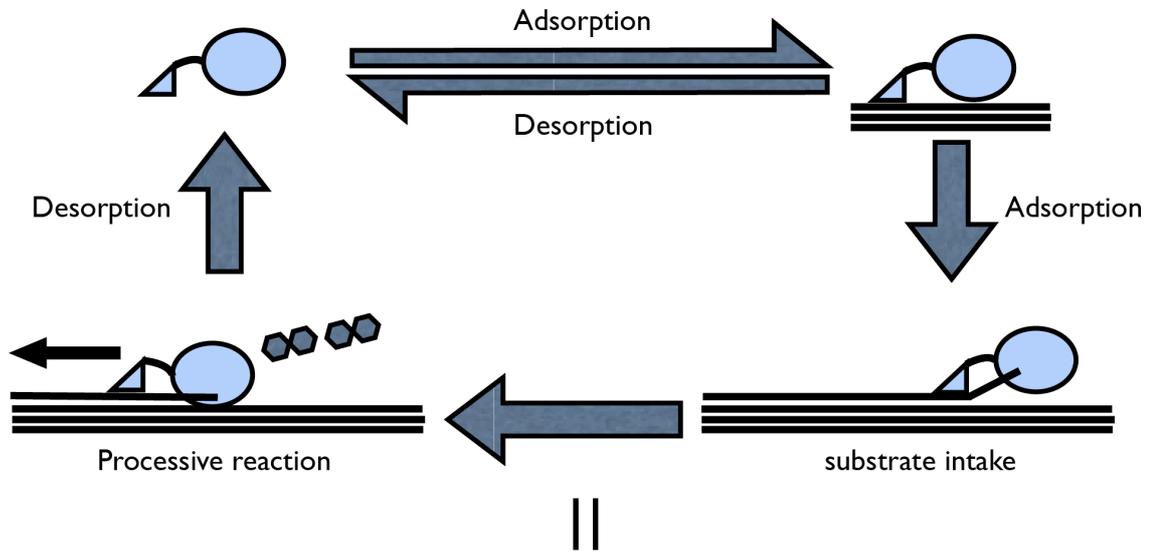
Direct comparison of the results from single molecule observation and biochemical activity measurement is difficult, because biochemical results contain also other effects.

For example,  $k_{on}$  of the CD,  $k_{on}$  and  $k_{off}$  of the CBM must affect. But the comparison of these results shows the importance of  $k_{cat}$ ,  $k_{off}$  and processivity for amorphous and crystalline cellulose degradation, because if one parameter is more important than the others, the tendency of hydrolysis rate determined by biochemical measurement should be similar to the tendency of the parameter. In PASC hydrolysis, it is shown that the order of hydrolysis activity seems to be the same order of moving-velocity and  $k_{off}$ . This result implies that the rate-limiting parameter of PASC degradation is  $k_{cat}$  or  $k_{off}$ . This hypothesis is supported by the report of less processive mutant of chitinase, which hydrolyze chitosan faster than wild type (121). In addition, because PASC has many reducing-ends, on which the CD can bind, binding on CBM or CD may not be a rate limiting step. In case of crystalline cellulose degradation, the tendency of activity is similar to that of processivity. This result implies that the binding of the catalytic domain is the rate-limiting step of crystalline cellulose degradation, because the processivity is not a single kinetic parameter. In the other words, the processivity is just a balance of two kinetic parameters ( $k_{cat}$  and  $k_{off}$ ) and indicates how many reactions will be catalyzed per a binding event of the catalytic domain. This hypothesis is supported by the synergy between processive enzymes and lytic polysaccharide monooxygenases, which make chain ends on the surface of crystalline substrate (56). In addition, the activity of *PcCel7D* is close to the activity of *TrCel7A*, not of *PcCel7C* when cellulose III<sub>I</sub> was substrate, but the activity of *TrCel7A* to crystalline cellulose I<sub>α</sub> was apparently higher than *PcCel7D*. These results also imply that the importance of processivity to degradation of crystalline cellulose III<sub>I</sub> was lower than that of cellulose I<sub>α</sub>, because cellulose III<sub>I</sub> has larger surface area and larger number of available chain ends than I<sub>α</sub>.

### 3.2.2 Validation of the processivity

It was experimentally shown that processivity defines the hydrolysis velocity of CBH on crystalline cellulose, but the relationship between processivity and hydrolysis velocity has not been theoretically explained. To clarify the importance of processivity for crystalline cellulose hydrolysis, the equation of production velocity was derived from a simple reaction model (Scheme 3-8). This model includes the adsorption/desorption by the CBD, chain end intake by the CD, processive reaction and dissociation of the CD. In this equation, if chain binding ( $k_{on}$ ) of the CD is enough faster than the dissociation rate constant ( $k_{off}$ ) of CD, the production velocity correlates with just  $k_{cat}$ . On the other hand, if  $k_{on}$  of CD is smaller than  $k_{off}$  of CD, the production velocity correlates with processivity ( $k_{cat}/k_{off}$ ),  $k_{on}$  of CD, and number of accessible chain ends. Unfortunately, the  $k_{on}$  values of the three CDs and the number of accessible chain ends of PASC and crystalline cellulose could not be determined yet. But in the case of crystalline cellulose, chain ends are apparently lesser accessible and number of chain ends per unit of mass or surface should be smaller than the values of amorphous cellulose. Thus, when the  $k_{on}$  of CD and the number of chain ends on crystalline cellulose are smaller, it is reasonable that processivity is more important for hydrolysis rate of crystalline cellulose than the hydrolysis rate of amorphous cellulose.

The roles of unique tryptophan residues with CBHs are increasing  $k_{on}$  of CD and processivity. These two factors both contribute to the hydrolysis velocity of less accessible cellulose. In addition, tunnel-like structure is also responsible to the processivity. Therefore, it can be said that unique structures in CBHs increase the hydrolysis efficiency to the less accessible crystalline cellulose.



$\alpha$  = number of chain ends / unit of cellulose

$$\frac{d[E]}{dt} = k_{off}^{CD} [E_{CD}C] + k_{off}^{CBM} [E_{CBM}C] - k_{on}^{CBM} [E][C]$$

When,  $k_{on}^{CD} \alpha \gg k_{off}^{CD}$

$$\frac{d[E_{CBM}C]}{dt} = k_{on}^{CBM} [E][C] - (k_{off}^{CBM} + k_{on}^{CD} \alpha) [E_{CBM}C]$$

$$\frac{d[P]}{dt} = k_{cat} [E_{CD}C] = \frac{k_{cat} \cdot E_0 \cdot [C]}{\frac{k_{off}^{CD} (k_{off}^{CBM} + k_{on}^{CD} \alpha)}{k_{on}^{CBM} \cdot k_{on}^{CD} \alpha} + [C]}$$

$$\frac{d[E_{CD}C]}{dt} = k_{on}^{CD} \alpha [E_{CBM}C] - k_{off}^{CD} [E_{CD}C]$$

When,  $k_{on}^{CD} \alpha \ll k_{off}^{CD}$

If in a steady state,

$$\frac{d[P]}{dt} = k_{cat} [E_{CD}C] = \frac{k_{cat} \cdot \frac{k_{on}^{CD} \alpha}{k_{off}^{CD} + k_{on}^{CD} \alpha} E_0 \cdot [C]}{\frac{k_{off}^{CD} (k_{off}^{CBM} + k_{on}^{CD} \alpha)}{k_{on}^{CBM} (k_{off}^{CD} + k_{on}^{CD} \alpha)} + [C]}$$

$$\frac{d[P]}{dt} = k_{cat} [E_{CD}C] = \frac{\frac{k_{cat}}{k_{off}^{CD}} \cdot k_{on}^{CD} \alpha \cdot E_0 \cdot [C]}{\frac{k_{off}^{CBM} + k_{on}^{CD} \alpha}{k_{on}^{CBM}} + [C]}$$

Scheme 3-8 Cellulose hydrolysis model of GH7 CBH

## **4. Conclusion**

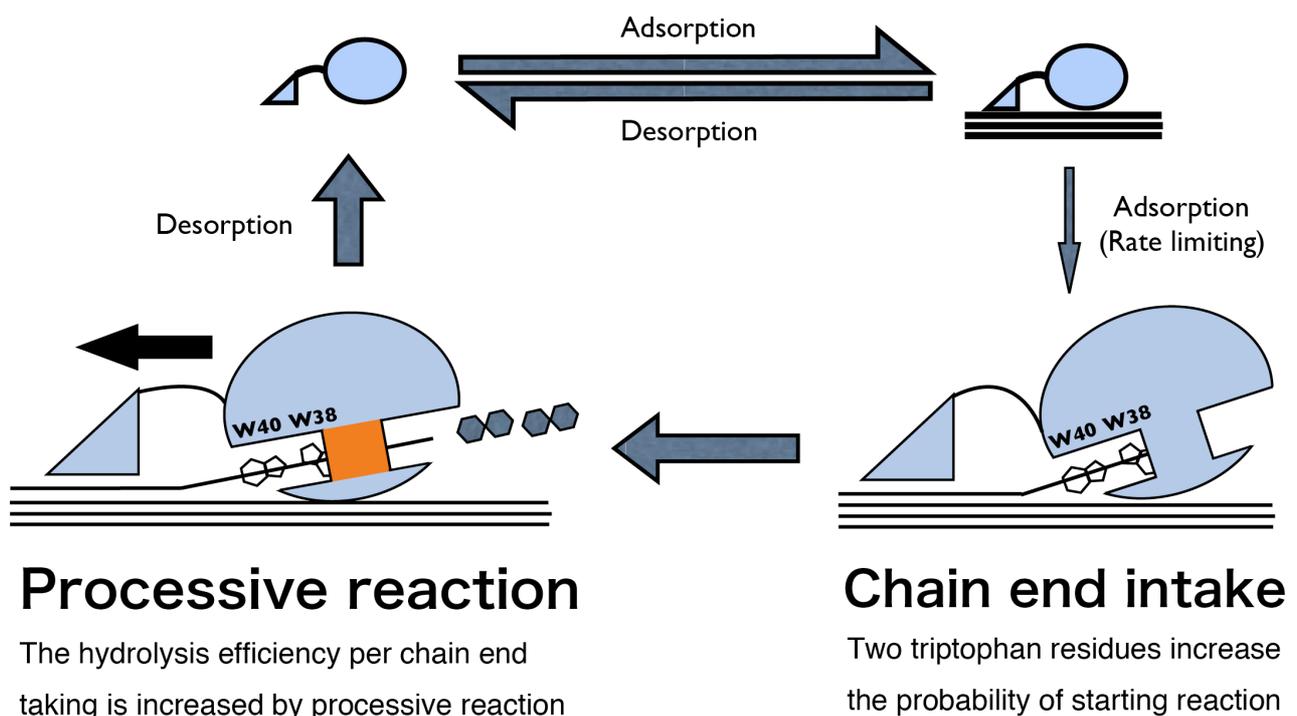
In this thesis, the aim was to clarify the efficient hydrolysis activity of GH7 CBH to the crystalline cellulose degradation focusing on the unique structure of CBHs. By the activity comparison among WT and tryptophan mutants of *TrCel7A*, the importance of the tryptophan residues uniquely conserved in CBHs for crystalline cellulose degradation was shown. Because the molecular chains are strongly packed in the crystalline cellulose, the chain ends are less accessible for cellulase. Therefore, the two tryptophan residues are specially needed to catch and initiate the hydrolysis. The results of molecular dynamics simulation showed that the tryptophan residue, which is located at the entrance of tunnel-like structure (subsite -7), play an important role in the intake of less accessible chain ends on crystalline cellulose surface. In addition, the results of cellotetraose hydrolysis indicate that the tryptophan residue at the subsite -4 keeps and sends the chain into the active center to start the processive hydrolysis of the molecular chain of cellulose.

Additionally, direct comparison of processive movements of the three different CBHs from *T. reesei* and *P. chrysosporium* by HS-AFM showed that the longer tunnel-like structure with more loop regions increases the processivity, but decreases the moving velocity. Comparison of the results of single-molecule analysis and biochemical analysis of enzyme activity for the crystalline and amorphous cellulose indicate that the enzyme with higher processivity has higher activity for strongly packed crystalline cellulose, even though the processive cycle is slow. This relationship between processivity and hydrolysis velocity of crystalline cellulose is also due to the less accessible structure of the crystalline cellulose. Because the probability of initiation of the processive reaction is very low, the hydrolysis efficiency per the chain intake is strongly related to the velocity of the whole cycle of crystalline cellulose degradation.

Therefore, based on all of the results in this thesis, it may be concluded that the structure of GH7 CBHs are designed to improve the chain end intake by the unique residues at the entrance of the tunnel and to increase the efficiency of hydrolysis per a

chain intake event by the processive reaction with sacrificing the hydrolysis efficiency to soluble or amorphous cellulose (Scheme 4-1). The unique structure of GH7 CBH seems to be tuned for the crystalline cellulose degradation. This conclusion implies that GH7 CBHs of filamentous fungi have been evolved to degrade the crystalline cellulose, which has been evolved by plants to provide strength to the plant cell wall.

Furthermore, the tryptophan residues at the entrance of tunnel are conserved among the CBHs from both ascomycete and basidiomycete, but the balance between processivity and velocity of processive cycle are different among them. The variety of CBH structures and the differences of balance between crystalline cellulose degradation efficiency and that of amorphous cellulose means the cellulose degradation strategy between ascomycete and basidiomycete are different. This result implies that the targeted substrate and kinds of other supporting enzymes are different between them. Therefore, efficiency of CBH should depend on the targeting substrate and on the used enzyme cocktail, and the possibility of improvement for the efficiency of biomass conversion was suggested.



Scheme 4-1 Crystalline cellulose degradation mechanism by GH7 cellobiohydrolases

# References

1. Atalla, R. H., and Vanderhart, D. L. (1984) Native Cellulose: A Composite of Two Distinct Crystalline Forms. *Science* **223**, 283-285
2. Yamamoto, H., and Horii, F. (1994) In Situ crystallization of bacterial cellulose I. Influences of polymeric additives, stirring and temperature on the formation celluloses Ia and Ib as revealed by cross polarization/magic angle spinning (CP/MAS)<sup>13</sup>C NMR spectroscopy. *Cellulose* **1**, 57-66
3. Wada, M., Okano, T., and Sugiyama, J. (1997) Synchrotron-radiated X-ray and neutron diffraction study of native cellulose. *Cellulose* **4**, 221-232
4. Hult, E.-L., Lason, P. T., and Iversen, T. (2002) A Comparative CP/MAS <sup>13</sup>C-NMR Study of the Supramolecular Structure of Polysaccharides in Sulphite and Kraft Pulps. *Holzforschung* **56**, 179-184
5. Belton, P. S., Tanner, S. F., Cartier, N., and Chanzy, H. (1989) High-resolution solid-state carbon-13 nuclear magnetic resonance spectroscopy of tunicin, an animal cellulose. *Macromolecules* **22**, 1615-1617
6. Yamamoto, H., Horii, F., and Odani, H. (1989) Structural changes of native cellulose crystals induced by annealing in aqueous alkaline and acidic solutions at high temperatures. *Macromolecules* **22**, 4130-4132
7. Sugiyama, J., Harada, H., Fujiyoshi, Y., and Uyeda, N. (1985) Lattice images from ultrathin sections of cellulose microfibrils in the cell wall of *Valonia macrophysa* Kütz. *Planta* **166**, 161-168
8. Doblin, M. S., Kurek, I., Jacob-Wilk, D., and Delmer, D. P. (2002) Cellulose Biosynthesis in Plants: from Genes to Rosettes. *Plant and Cell Physiology* **43**, 1407-1420
9. Brown, R. M., and Montezinos, D. (1976) Cellulose microfibrils: visualization of biosynthetic and orienting complexes in association with the plasma membrane. *Proceedings of the National Academy of Sciences* **73**, 143-147
10. Herth, W. (1983) Arrays of plasma-membrane "rosettes" involved in cellulose microfibril formation of Spirogyra. *Planta* **159**, 347-356
11. ROSS, P., MAYER, R., and BENZIMAN, M. (1991) Cellulose biosynthesis and function in bacteria. *Microbiological Reviews* **55**, 35-58
12. Watanabe, S., Kimura, K., and Akahori, T. (1968) A Study of Mercerization of Native Cellulose by X-Ray Method. *Bulletin of the Faculty of Engineering, Hokkaido University* **47**, 121-130
13. Manjunath, B. R., and Venkataraman, A. (1980) Fibrillar aggregation in cotton cellulose subjected to multiple swelling treatments with alkali. *Journal of Polymer Science: Polymer Chemistry Edition* **18**, 1407-1424
14. Sugiyama, J., and Horikawa, Y. (2008) The Cellulose Microfibril and Its Polymorphs. *Mokuzai Gakkaishi* **54**, 49-57
15. Barry, A. J., Peterson, F. C., and King, A. J. (1936) X-Ray Studies of Reactions of Cellulose in Non-Aqueous Systems. I. Interaction of Cellulose and Liquid Ammonia. *Journal of the American Chemical Society* **58**, 333-337
16. Yatsu, L. Y., Calamari, T. A., and Benerito, R. R. (1986) Conversion of Cellulose I to Stable Cellulose III. *Textile Research Journal* **56**, 419-424
17. Wada, M., Heux, L., Nishiyama, Y., and Langan, P. (2009) The structure of the complex of cellulose I with ethylenediamine by X-ray crystallography and cross-polarization/magic angle spinning <sup>13</sup>C nuclear magnetic resonance. *Cellulose* **16**, 943-957

18. Wada, M. (2001) In Situ Observation of the Crystalline Transformation from Cellulose III to I $\beta$ . *Macromolecules* **34**, 3271-3275
19. Buleon, A., and Chanzy, H. (1980) Single crystals of cellulose IVII: Preparation and properties. *Journal of Polymer Science: Polymer Chemistry Edition* **18**, 1209-1217
20. Wada, M., Heux, L., and Sugiyama, J. (2004) Polymorphism of Cellulose I Family: Reinvestigation of Cellulose IVI. *Biomacromolecules* **5**, 1385-1391
21. Nishiyama, Y., Sugiyama, J., Chanzy, H., and Langan, P. (2003) Crystal Structure and Hydrogen Bonding System in Cellulose Ia from Synchrotron X-ray and Neutron Fiber Diffraction. *Journal of the American Chemical Society* **125**, 14330-14306
22. Nishiyama, Y., Langan, P., and Chanzy, H. (2002) Crystal Structure and Hydrogen-Bonding System in Cellulose I $\beta$  from Synchrotron X-ray and Neutron Fiber Diffraction. *Journal of the American Chemical Society* **124**, 9074-9082
23. Langan, P., Nishiyama, Y., and Chanzy, H. (1999) A Revised Structure and Hydrogen-Bonding System in Cellulose II from a Neutron Fiber Diffraction Analysis. *Journal of the American Chemical Society* **121**, 9940-9946
24. Langan, P., Nishiyama, Y., and Chanzy, H. (2001) X-ray Structure of Mercerized Cellulose II at 1 Å Resolution. *Biomacromolecules* **2**, 410-416
25. Wada, M., Chanzy, H., Nishiyama, Y., and Langan, P. (2004) Cellulose III I Crystal Structure and Hydrogen Bonding by Synchrotron X-ray and Neutron Fiber Diffraction. *Macromolecules* **37**, 8548-8555
26. Pringsheim, H. (1912) Utber der fermentativen Abbau der Zellulose. *Zeitschrift fur physiologische Chemie* **78**, 266-291
27. Bradley, L. A., and Rettger, L. F. (1927) Studies on aerobic bacteria commonly concerned in the decomposition of cellulose. *Journal of Bacteriology* **13**, 321
28. Reese, E. T., Siu, R. G., and Levinson, H. S. (1950) The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *Journal of Bacteriology* **59**, 485
29. Gilligan, W., and Reese, E. T. (1954) Evidence for multiple components in microbial cellulases. *Canadian Journal of Microbiology* **1**, 90-107
30. Selby, K., and Maitland, C. C. (1967) *The cellulase of Trichoderma viride*, *Biochemical Journal* **104**, 716-724
31. Li, L. H., Flora, R. M., and King, K. W. (1965) Individual roles of cellulase components derived from *Trichoderma viride*. *Archives of Biochemistry and Biophysics* **111**, 439-447
32. Wood, T. M. and Mccrae, S. I. (1972) The purification and properties of the C1 component of *Trichoderma koningii* cellulase. *Biochemical Journal* **128**, 1183-1192
33. Berghem, L. E. R., and Pettersson, L. G. (1973) The mechanism of enzymatic cellulose degradation. Purification of a cellulolytic enzyme from *Trichoderma viride* active on highly ordered cellulose. *European journal of biochemistry / FEBS* **37**, 21-30
34. Eriksson, K.-E. (1969) *Cellulases and Their Applications*, American Chemical Society Publications, Washington, D.C
35. Eriksson, K.-E., and Pettersson, B. (1972) *Biodeterioration of Materials*, Applied Sciences Publishers Ltd, London.
36. Streamer, M., Eriksson, K.-E., and Pettersson, B. (1975) Extracellular enzyme system utilized by the fungus *Sporotrichum pulverulentum* (*Chrysosporium lignorum*) for the breakdown of cellulose. functional characterization of five endo-1,4-beta-glucanases and one exo-1,4-beta-glucanase. *European journal of biochemistry / FEBS* **59**, 607-613
37. Wood, T. M. and Mccrae, S. I. (1978) The cellulase of *Trichoderma koningii*. Purification and properties of some endoglucanase components with special reference to their action on cellulose when acting alone and in synergism with the

- cellobiohydrolase. *Biochemical Journal* **171**, 61
38. Gritzai, M., and Brown, R. D. (1979) The cellulase system of *Trichoderma* relationships between purified extracellular enzymes from induced or cellulose-grown cells. *Hydrolysis of Cellulose: Mechanisms of Enzymatic and Acid Catalysis* American Chemical Society, Washington, D. C. Chapter 12, pp 237-260
  39. Beldman, G., Leeuwen, M. F., Rombouts, F. M., and Voragen, F. G. J. (1985) The cellulase of *Trichoderma viride*. Purification, characterization and comparison of all detectable endoglucanases, exoglucanases and beta-glucosidases. *European journal of biochemistry / FEBS* **146**, 301-308
  40. Fägerstam, L. G., and Pettersson, L. G. (1980) The 1.4- $\beta$ -glucan cellobiohydrolases of *Trichoderma reesei* QM 9414. *FEBS letters* **119**, 97-100
  41. Teeri, T., Salovuori, I., and Knowles, J. (1983) The Molecular Cloning of the Major Cellulase Gene from *Trichoderma reesei*. *Bio-Technology* **1**, 696-699
  42. Penttilä, M., Lehtovaara, P., Nevalainen, H., Bhikhabhai, R., and Knowles, J. (1986) Homology between cellulase genes of *Trichoderma reesei*: complete nucleotide sequence of the endoglucanase I gene. *Gene* **45**, 253-263
  43. Teeri, T. T., Lehtovaara, P., Kauppinen, S., Salovuori, I., and Knowles, J. (1987) Homologous domains in *Trichoderma reesei* cellulolytic enzymes: Gene sequence and expression of cellobiohydrolase II. *Gene* **51**, 43-52
  44. Panttilä, M. E., André, L., Saloheimo, M., Lehtovaara, P., and Knowles, J. K. C. (1987) Expression of two *Trichoderma reesei* endoglucanases in the yeast *Saccharomyces cerevisiae*. *Yeast* **3**, 175-185
  45. Arai, M., Sakamoto, R., and Murao, S. (1989) Different action by two avicelases from *Aspergillus aculeatus*. *Agricultural And Biological Chemistry* **53**, 1411-1412
  46. Claeysens, M., Tilbeurgh, H. V., Tomme, P., Wood, T. M., and Mcrae, S. I. (1989) Fungal cellulase systems. Comparison of the specificities of the cellobiohydrolases isolated from *Penicillium pinophilum* and *Trichoderma reesei*. *Biochemical Journal* **261**, 819-825
  47. Vršanská, M., and Biely, P. (1992) The cellobiohydrolase I from *Trichoderma reesei* QM 9414: action on cello-oligosaccharides. *Carbohydrate research* **227**, 19-27
  48. Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K., and Jones, T. A. (1990) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science (New York, NY)* **249**, 380-386
  49. Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K., Teeri, T. T., and Jones, T. A. (1994) The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science (New York, NY)* **265**, 524-528
  50. Spezio, M., Wilson, D. B., and Karplus, P. A. (1993) Crystal structure of the catalytic domain of a thermophilic endocellulase. *Biochemistry* **32**, 9906-9916
  51. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859
  52. Teeri, T. T. (1997) Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends In Biotechnology* **15**, 160-167
  53. Henrissat, B., Claeysens, M., Tomme, P., LEMESLE, L., and MORNON, J. P. (1989) Cellulase families revealed by hydrophobic cluster analysis. *Gene* **81**, 83-95
  54. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical journal* **280**, 309-316
  55. Vaaje-Kolstad, G., Westereng, B., Horn, S., Liu, Z., Zhai, H., Sørlie, M., and Eijsink, V. (2010) An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science (New York, NY)* **330**, 219
  56. Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E., and Sweeney, M. D. (2011) Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose

- Dehydrogenase and Glycoside Hydrolase 61. *Applied and Environmental Microbiology* **77**, 7007-7015
57. Quinlan, R. J., Sweeney, M. D., Leggio, L. L., Otten, H., Poulsen, J.-C. N., Johansen, K. S., Krogh, K. B., Jørgensen, C. I., Tovborg, M., and Anthonsen, A. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences* **108**, 15079-15084
  58. Horn, S., Vaaje-Kolstad, G., Westereng, B., and Eijsink, V. (2012) Novel enzymes for the degradation of cellulose. *Biotechnology For Biofuels* **5**, 45-45
  59. Levasseur, A., Drula, E., Lombard, V., Coutinho, P., and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology For Biofuels* **6**, 41-41
  60. Enzyme Nomenclature Supplement 7. (2001) in *Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB)*
  61. Enzymes, Congress of International Union of Biochemistry. (1961) Report of the Commission on Enzymes of the International Union of Biochemistry 1961. in *I.U.B. symposium series*
  62. Enzymes, Congress of International Union of Biochemistry. (1976) Enzyme nomenclature: Recommendations (1972) of the International Union of Pure and Applied Chemistry and the International Union of Biochemistry: Supplement I: Corrections & additions (1975). *Biochimica et Biophysica Acta (BBA) - Enzymology* **429**, 1-45
  63. Enzyme Nomenclature Supplement 17. (2011) in *Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB)*
  64. Nagel, B., Dellweg, H., and Gierasch, L. M. (1992) Glossary for chemists of terms used in biotechnology (IUPAC Recommendations 1992) *International Union of Pure and Applied Chemistry* **64**, 143-168
  65. Bambara, R. A., Uyemura, D., and Lehman, I. R. (1976) On the processive mechanism of *Escherichia coli* DNA polymerase I. Delayed initiation of polymerization. *Journal of Biochemistry* **251**, 4090-4094
  66. Kondo, H., Nakatani, H., Matsuno, R., and Hiromi, K. (1980) Product distribution in amylase-catalyzed hydrolysis of amylose comparison of experimental results with theoretical predictions. *Journal of Biochemistry* **87**, 1053-1070
  67. Koshland, D. E. (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biological Reviews* **28**, 416-436
  68. McCarter, J. D., and Stephen Withers, G. (1994) Mechanisms of enzymatic glycoside hydrolysis. *Current Opinion in Structural Biology* **4**, 885-892
  69. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Körner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) The pK<sub>a</sub> of the general acid/base carboxyl group of a glycosidase cycles during catalysis: A <sup>13</sup>C-NMR study of *Bacillus circulans* xylanase. *Biochemistry* **35**, 9958-9966
  70. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* **37**, D233-D238
  71. B, H., and A, B. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal* **316**, 695–696
  72. Eriksson, K.-E., and Pettersson, B. (1975) Extracellular enzyme system utilized by the fungus *Sporotrichum pulverulentum* (*Chrysosporium lignorum*) for the breakdown of cellulose. 3. Purification and physico-chemical characterization of an exo-1,4-beta-glucanase. *European journal of biochemistry / FEBS* **51**, 213-218
  73. Igarashi, K., Koivula, A., Wada, M., Kimura, S., Penttila, M., and Samejima, M.

- (2009) High speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *The Journal of Biological Chemistry* **284**, 36186-36190
74. Payne, C. M., Resch, M. G., and Chen, L. (2013) Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose. *Proceedings of the National Academy of Sciences* **110**, 14646-14651
  75. Divne, C., Ståhlberg, J., Teeri, T. T., and Jones, T. A. (1998) High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *Journal of Molecular Biology* **275**, 309-325
  76. Munoz, I., Ubhayasekera, W., Henriksson, H., Szabo, I., Pettersson, G., Johansson, G., Mowbray, S., and Stahlberg, J. (2001) Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Å resolution and homology models of the isozymes. *Journal of Molecular Biology* **314**, 1097-1111
  77. Knowles, J. K. C., Lentovaara, P. i., Murray, M., and Sinnott, M. L. (1988) Stereochemical course of the action of the cellobioside hydrolases I and II of *Trichoderma reesei*. *Journal of the Chemical Society-Chemical Communications*, 1401-1402
  78. Stahlberg, J., Divne, C., Koivula, A., Piens, K., Claeysens, M., Teeri, T. T., and Jones, T. A. (1996) Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from *Trichoderma reesei*. *Journal of Molecular Biology* **264**, 337-349
  79. MacKenzie, L., Sulzenbacher, G., Divne, C., Jones, T., Woldik, H., Schulein, M., Withers, S., and Davies, G. (1998) Crystal structure of the family 7 endoglucanase I (Cel7B) from *Humicola insolens* at 2.2 Å resolution and identification of the catalytic nucleophile by trapping of the covalent glycosyl-enzyme intermediate. *The Biochemical journal* **335**, 409-416
  80. Klarskov, K., Piens, K., Stahlberg, J., Høj, P. B., Van Beeumen, J., and Claeysens, M. (1997) Cellobiohydrolase I from *Trichoderma reesei*: Identification of an active-site nucleophile and additional information on sequence including the glycosylation pattern of the core protein. *Carbohydrate research* **304**, 143-154
  81. Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., and Gronenborn, A. M. (1989) Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry* **28**, 7241-7257
  82. Mattinen, M.-L., Linder, M., Teleman, A., and Annala, A. (1997) Interaction between cellohexaose and cellulose binding domains from *Trichoderma reesei* cellulases. *FEBS letters* **407**, 291-296
  83. Stals, I., Sandra, K., Geysens, S., Contreras, R., Van Beeumen, J., and Claeysens, M. (2004) Factors influencing glycosylation of *Trichoderma reesei* cellulases. I: Postsecretorial changes of the O- and N-glycosylation pattern of Cel7A. *Glycobiology* **14**, 713-724
  84. Divne, C., Stahlberg, J., Teeri, T. T., and Jones, T. A. (1998) High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *Journal of Molecular Biology* **275**, 309-325
  85. Rose, I. A., Hanson, K. R., Wilkinson, K. D., and Wimmer, M. J. (1980) A suggestion for naming faces of ring compounds. *Proceedings of the National Academy of Sciences* **77**, 2439
  86. Koivula, A., Kinnari, T., Harjunpää, V., Ruohonen, L., Teleman, A., Drakenberg, T.,

- Rouvinen, J., Jones, T. A., and Teeri, T. T. (1998) Tryptophan 272: an essential determinant of crystalline cellulose degradation by *Trichoderma reesei* cellobiohydrolase Cel6A. *FEBS Letters* **429**, 341-346
87. Maguire, R. J. (1977) Kinetics of the hydrolysis of cellulose by beta-1,4-glucan cellobiohydrolase of *Trichoderma viride*. *Can J Biochem.* **6**, 644-650
88. Chanzy, H., Henrissat, B., Vuong, R., and Schulein, M. (1983) The action of 1,4- $\beta$ -D-glucan cellobiohydrolase on Valonia cellulose microcrystals. *FEBS letters* **153**, 113-118
89. Schmuck, M., Pilz, I., Hayn, M., and Esterbauer, H. (1986) Investigation of cellobiohydrolase from *Trichoderma reesei* by small angle X-ray scattering. *Biotechnology letters* **8**, 397-402
90. Van Tilbeurgh, H., Tomme, P., Claeysens, M., Bhikhabhai, R., and Pettersson, G. (1986) Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*. *FEBS letters* **204**, 223-227
91. Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., Claeysens, M. (1988) Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. *European Journal of Biochemistry* **170**, 575-581
92. Carrard, G., and Linder, M. (1999) Widely different off rates of two closely related cellulose-binding domains from *Trichoderma reesei*. *European journal of biochemistry / FEBS* **262**, 637-643
93. Palonen, H., Tenkanen, M., and Linder, M. (1999) Dynamic interaction of *Trichoderma reesei* cellobiohydrolases Cel6A and Cel7A and cellulose at equilibrium and during hydrolysis. *Applied and Environmental Microbiology* **65**, 5229-5233
94. Stahlberg, J., Johansson, G., and Pettersson, G. (1991) A new model for Enzymatic hydrolysis of cellulose based on the two-domain structure of cellobiohydrolase I. *Bio-Technology* **9**, 286-290
95. Sugimoto, N., Igarashi, K., Wada, M., and Samejima, M. (2012) Adsorption characteristics of fungal family 1 cellulose-binding domain from *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose: negative cooperative adsorption via a steric exclusion effect. *Langmuir* **28**, 14323-14329
96. Várnai, A., Siika-aho, M., and Viikari, L. (2013) Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnol Biofuels* **6**, 30-41
97. Stahlberg, J., Johansson, G., and Pettersson, G. (1993) *Trichoderma reesei* has no true exo-cellulase: all intact and truncated cellulases produce new reducing end groups on cellulose. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1157**, 107-113
98. Armand, S., Drouillard, S., Schulein, M., Henrissat, B., and Driguez, H. (1997) A bifunctionalized fluorogenic tetrasaccharide as a substrate to study cellulases. *Journal Of Biochemistry* **272**, 2709-2713
99. Kurasin, M., and Valjamae, P. (2011) Processivity of Cellobiohydrolases is limited by the substrate. *Journal of Biological Chemistry* **286**, 169-177
100. Imai, T., Boisset, C., Samejima, M., Igarashi, K., and Sugiyama, J. (1998) Unidirectional processive action of cellobiohydrolase Cel7A on Valonia cellulose microcrystals. *FEBS letters* **432**, 113-116
101. Valjamae, P., Sild, V., Pettersson, G., and Johansson, G. (1998) The initial kinetics of hydrolysis by cellobiohydrolases I and II is consistent with a cellulose surface - erosion model. *European Journal of Biochemistry / FEBS* **253**, 469-475
102. Medve, J., Karlsson, J., Lee, D., and Tjerneld, F. (1998) Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the

- enzymes. *Biotechnology and Bioengineering* **59**, 621-634
103. von Ossowski, I., Stahlberg, J., Koivula, A., Piens, K., Becker, D., Boer, H., Harle, R., Harris, M., Divne, C., Mahdi, S., Zhao, Y., Driguez, H., Claeysens, M., Sinnott, M. L., and Teeri, T. T. (2003) Engineering the Exo-loop of *Trichoderma reesei* Cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium* Cel7D. *Journal of Molecular Biology* **333**, 817-829
  104. Kipper, K., Valjamae, P., and Johansson, G. (2005) Processive action of cellobiohydrolase Cel7A from *Trichoderma reesei* is revealed as 'burst' kinetics on fluorescent polymeric model substrates. *Biochemical Journal* **385**, 527
  105. Mulakala, C., and Reilly, P. J. (2005) *Hypocrea jecorina* (*Trichoderma reesei*) Cel7A as a molecular machine: A docking study. *Proteins-Structure Function And Bioinformatics* **60**, 598-598-605-605
  106. Jalak, J., Kurasin, M., Teugjas, H., and Valjamae, P. (2012) Endo-exo synergism in cellulose hydrolysis revisited. *Journal of Biological Chemistry* **287**, 28802-28815
  107. Hsu, T.-A., Gong, C.-S., and Tsao, G. T. (1980) Kinetic studies of cellodextrins hydrolyses by exocellulase from *Trichoderma reesei*. *Biotechnology and Bioengineering* **22**, 2305-2320
  108. Fan, L. T., Lee, Y.-H., and Beardmore, D. H. (1980) Mechanism of the enzymatic hydrolysis of cellulose: Effects of major structural features of cellulose on enzymatic hydrolysis. *Biotechnology and Bioengineering* **22**, 177-199
  109. Huang, A. A. (1975) Kinetic studies on insoluble cellulose-cellulase system. *Biotechnology and Bioengineering* **17**, 1421-1433
  110. Beldman, G., Voragen, A. G. J., Rombouts, F. M., Searle-van Leeuwen, M. F., and Pilnik, W. (1987) Adsorption and kinetic behavior of purified endoglucanases and exoglucanases from *Trichoderma viride*. *Biotechnol. Bioeng.* **30**, 251-257
  111. Igarashi, K., Wada, M., Hori, R., and Samejima, M. (2006) Surface density of cellobiohydrolase on crystalline celluloses. A critical parameter to evaluate enzymatic kinetics at a solid-liquid interface. *FEBS journal* **273**, 2869-2878
  112. Igarashi, K., Wada, M., and Samejima, M. (2007) Activation of crystalline cellulose to cellulose III(I) results in efficient hydrolysis by cellobiohydrolase. *FEBS journal* **274**, 1785-1792
  113. Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Okamoto, T., Penttilä, M., Ando, T., and Samejima, M. (2011) Traffic Jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science (New York, NY)* **333**, 1279-1282
  114. Jalak, J., and Valjamae, P. (2010) Mechanism of initial rapid rate retardation in cellobiohydrolase catalyzed cellulose hydrolysis. *Biotechnology and Bioengineering* **106**, 871-883
  115. Case, D. A., Darden, T. A., Cheatham, T. E., 3rd, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010). *AMBER 11, University of California, San Francisco.*
  116. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics* **81**, 3684-3690
  117. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics* **98**, 10089-10092
  118. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics.

*Journal of Molecular Graphics* **14**, 33-38

119. Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Penttilä, M., Ando, T., and Samejima, M. (2012) Chapter nine – Visualization of cellobiohydrolase I from *Trichoderma reesei* moving on crystalline cellulose using high-speed atomic force microscopy. *Methods in enzymology* **510**, 169-182
120. Payne, C. M., Jiang, W., Shirts, M. R., Himmel, M. E., Crowley, M. F., and Beckham, G. T. (2013) Glycoside Hydrolase Processivity Is Directly Related to Oligosaccharide Binding Free Energy. *Journal of the American Chemical Society* **135**, 18831-18839
121. Horn, S. J., Sikorski, P., Cederkvist, J. B., Vaaje-Kolstad, G., Sørlie, M., Synstad, B., Vriand, G., Vårum, K. M., and Eijsink, V. G. H. (2006) Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proceedings of the National Academy of Sciences* **103**, 18089-18094



## 謝辞 / Acknowledgement

本研究のテーマは学部生として森林化学研究室に配属されてから6年間、一貫して取り組んできた物です。まずこの研究を始める機会を頂いた事について、鮫島正浩教授及び五十嵐圭日子准教授に非常に感謝しております。そしてこの研究がまとまったのもお二人のご指導有ったの賜物です。今振り返ってみるとお二人には、適宜良いタイミングで実験を進める機会と御指導をいただいていたと気づかされます。心より御礼申し上げます。

また森林化学研究室での6年間の間において、寺田珠実助教には研究室の運営などお世話になりました。また石田卓也特任助教には学部生の頃から酵素の構造の見方や解析の仕方に始まり、研究全般について様々なご指導を頂きました。現東レ株式会社塚田剛士研究員には、森林化学研究室に所属し日本学術振興会特別研究員でいらっしゃったときに、本研究に用いた*TrCel7A*の変異体をフィンランドのVTTで作成していただき、また投稿論文用のデータの取得もしていただきました。現産業総合研究所石黒真希研究員及び現東京文化財研究所和田朋子研究員には遺伝子操作技術や配列解析技術など教えていただきました。また現理化学研究所鈴木一史研究員、堀千明研究員には*Phanerochaete chrysosporium*の培養やcDNAの合成方法など教えていただきました。現新潟大学杉本直久特任助教にはGH7に付属しているCBM1の吸着挙動について教えていただいたり、また研究全般の面白さについて教えていただきました。また平石正男先輩には酵素の精製及びHPLCでの酵素活性の測定法などこの研究のベースとなる技術を教えていただきました。

また東京大学大学院農学生命科学研究科 生物材料科学専攻の木材化学研究室の松本雄二教授、生物素材化学研究室の和田正昌准教授、また応用生命工学専攻の酵素学研究室の伏信進矢教授に副査をしていただきました。東京工業大学古田忠臣助教には*TrCel7A*の分子動力学シミュレーションを行っていただきました。そして金沢大学安藤敏夫教授、内橋貴之准教授及び渡邊大輝様には高速AFMでのセルラーゼ観察を行っていただきました。皆様には心より感謝しております。 I was also supported by researchers Anu Koivula and Saner Auer of VTT Technical Research Centre of Finland. I'm deeply grateful for their producing *TrCel7A* mutants. And I appreciate very much the the support of English writing from Jenni Rahikainen.

そして9年間も大学に通わせてくださった両親には本当に感謝しております。多分大学は4年間だと思っていたでしょうが、まさかその後大学院に5年間も通うとは思っていません。本当にご迷惑おかけしました。また妻のご両親には、学生結婚を許していただいた事に加え、出張の度に子供の世話までしていただきありがとうございます。いつもご心配おかけしてすみません。そして妻の道子と長男の瞭太にはいつも支えていただいています。二人とも本当にありがとうございます。

平成25年 3月 3日

# List of publications

1. “The Tryptophan Residue at the Active Site Tunnel Entrance of *Trichoderma reesei* Cellobiohydrolase Cel7A Is Important for Initiation of Degradation of Crystalline Cellulose ”

The Journal of Biological Chemistry (2013) Vol.288, pp.13503-13510

2. “Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose ”

Journal of the American Chemical Society (accepted)