J. Appl. Glycosci., **71**, 63−72 (2024)
doi:10.5458/jag.jag.JAG-2023_0018

Regular Paper

Thermotolerance Mechanism of Fungal GH6 Cellobiohydrolase. Part II. Structural Analysis of Thermotolerant Mutant from the Basidiomycete *Phanerochaete chrysosporium*

(Received December 18, 2023; Accepted April 4, 2024)

Sora Yamaguchi,[1] Naoki Sunagawa,[1] Masahiro Samejima,[1,*] and Kiyohiko Igarashi[1,†]

[1] *Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo* (*1−1−1 Yayoi, Bunkyo-ku, Tokyo 113−8657, Japan*)

Abstract: Glycoside hydrolase family 6 cellobiohydrolase (GH6 CBH) is a group of cellulases capable of hydrolyzing crystalline cellulose. However, the synergistic reaction of GH6 CBH with other cellulases is hindered by its relatively low thermotolerance. We previously obtained a thermotolerant double mutant, C240S/C393S, of GH6 CBH from the basidiomycete *Phanerochaete chrysosporium* (*Pc*Cel6A) by replacing the two free cysteine (Cys) residues, C240 and C393, with serine (Yamaguchi *et al*., J Appl Glycosci. 2020; 67;79-86). In the accompanying paper (Part I; Yamaguchi *et al*., J Appl Glycosci. 2024; 71: 55−62), we measured the temperature dependence of the activity and folding of C240S/C393S and its single mutants, C240S and C393S, and found that replacement of C393 was the major contributor to the increased thermotolerance of C240S/C393S. Here, in order to investigate the mechanism involved, we crystallized the wild-type and the mutant enzymes and compared their X-ray crystal structures. The overall structures of the wild-type and the three mutant enzymes were similar. However, C240S/C393S had the lowest relative *B*-factor at both the N-terminal loop (residues 172−177) and the C-terminal loop (residues 390−425). This result suggests that reduced structural fluctuation of the substrate-enclosing loops, possibly due to stronger hydrogen bonding involving C393, could account for the increased thermotolerance of C240S/C393S.

Key words: Cellulose, cellobiohydrolase, GH6, thermotolerance, free cysteine, X-ray crystal structure

INTRODUCTION

Glycoside hydrolase family 6 (GH6) in the Carbohydrate Active enZymes database (CAZy) [1] is a group of cellulases mainly secreted by microorganisms such as fungi and bacteria that live on lignocellulose. There are two types, with different modes of hydrolysis. Endoglucanases (EGs), classified as EC 3.2.1.4, randomly degrade amorphous regions of cellulose, while cellobiohydrolases (CBHs), classified as EC 3.2.1.91, hydrolyze crystalline regions of cellulose from the non-reducing end [2]. This distinctive feature of CBHs is mainly attributed to the carbohydrate-binding module (CBM) and the presence of a relatively long substrate-enclosing loop compared to EGs. A CBM is connected to the N-terminus of the GH6 CBH catalytic domain via a flexible linker [3] and serves to bind the enzyme to the surface of crystalline cellulose [4]. Then, hydrolysis is initiated by the introduction of a single cellulose chain into the active site with the aid of the N- and C-terminal loops.

Since the degradation of the recalcitrant crystalline region is likely to be the rate-limiting step, the efficiency of CBH is critical in industrial cellulose saccharification, which aims to produce soluble sugars as feedstocks for biochemicals and biofuels. In addition to GH6 CBH, there is another group of CBHs, glycoside hydrolase family 7 (GH7), which is classified as EC 3.2.1.176. While GH6 CBHs exhibit endo-initiation as well as exo-initiation [2] and are less processive [5], GH7 CBHs exhibit exclusively exo-initiation [2] with high processivity [5]. These two contrasting CBH families synergistically degrade crystalline cellulose [5−7], because GH7 CBHs start hydrolysis from the reducing end, in contrast to GH6 CBHs [8]. It is desirable to perform cellulose degradation at high temperature in an industrial context in order to increase the cellulose degradation rate, but GH6 CBHs are less thermotolerant than GH7 CBHs, so this is problematic [9, 10]. Hence, many attempts have been made to increase the activity of GH6 CBH at high temperatures [11−13]. Notably, Arnold's lab reported that an engineered chimeric GH6 CBH could be made thermotolerant by substituting free (non-disulfide-forming) Cys residues [14, 15]. Based on that work, we designed a mutant of GH6 CBH from the basidiomycete *Phanerochaete chrysosporium* (*Pc*Cel6A) in which the two free Cys residues were replaced with serine, C240S and C393S, in order to explore the thermotolerance mechanism of GH6 CBH [16]. This double mutant, C240S/C393S, was confirmed to be thermotolerant,

†Corresponding author (Tel. +81−3−5841−5255, E-mail: aquarius@mail.ecc.u-tokyo.ac.jp, ORCID ID: 0000-0001-5152-7177)
Abbreviations: CAZy, Carbohydrate-Active enZymes database; CBH, cellobiohydrolase; CBM, carbohydrate-binding module; EG, endoglucanase; GH6, glycoside hydrolase family 6; GH7, glycoside hydrolase family 7; PEG, polyethylene glycol; WT, wild type.
*Present address: Masahiro Samejima, Faculty of Engineering, Shinshu University (4−17−1 Wakasato, Nagano 380−8553, Japan)

and the activity and thermal denaturation of C240S/C393S itself and the single mutants C240S and C393S at different temperatures were investigated in the accompanying paper [17]. The results indicated that the substitution at C393 was the major contributor to the increased activity at elevated temperature compared to the wild type (WT), as well as to the suppression of heat-induced loss of secondary structure.

In this study, we aimed to understand the mechanism of the increased thermotolerance by determining and comparing the X-ray crystal structures of the wild-type, double-mutant C240S/C393S, and the two single mutants, C240S and C393S. Our results suggest that the increased thermotolerance is predominantly due to reduced structural fluctuation of the substrate-enclosing loops in the catalytic domain.

MATERIALS AND METHODS

Materials. Sodium chloride, sodium acetate trihydrate, acetic acid, 2-amino-2-hydroxymethyl-1,3-propanediol (tris), and 6 mol/L hydrochloric acid (FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) and 50 % (w/v) polyethylene glycol (PEG) 4000 (NeXtal Biotechnologies, Holland, OH, USA) were used for the crystallization solution.

Protein preparation. The catalytic domains (residues 82−439) of *Pc*Cel6A C240S, C393S, and C240S/C393S were prepared, expressed, and purified as described in our accompanying paper (part I).

Crystallization. The protein was concentrated to 17.0−30.4 mg/mL with a Vivaspin™ 500-5K (Sartorius, Göttingen, Germany) and filtered with Nanosep MF Centrifugal Devices (Bio-Inert® Membrane 0.45 μm; Pall Corporation, Port Washington, NY, USA). As previously described [18], the protein was crystallized at 20 ˚C by the sitting drop vapor diffusion method [19, 20]. The reservoir solution comprised 50 mM sodium acetate (pH 4.5 or 5.0), 10 or 20 % (w/v) PEG 4000, and 0 or 140 mM NaCl. The initial concentration of droplets before diffusion was about 8.5−15.2 mg/mL protein, 10 or 80 mM NaCl, 5 or 10 % (w/v) PEG 4000, 25 mM sodium acetate (pH 4.5 or 5.0), and 10 mM Tris-HCl (pH 8.0), as summarized in Table S1 (see J. Appl. Glycosci. Web site). Each crystal was immersed in a cryoprotectant solution containing 40 % (w/v) PEG 4000 in the reservoir solution and frozen in liquid nitrogen.

Structure analysis. X-ray diffraction data were obtained at 95 K using synchrotron radiation at the BL-5A X-ray beamline of the Photon Factory in the High Energy Accelerator Research Organization, Tsukuba, Japan. The diffraction data were automatically processed, scaled, and merged using XDS [21], POINTLESS [22], and AIMLESS [23] integrated into PReMo [24]. Molecular replacement was performed with PHENIX (v.1.19.2 and 1.20.1) [25] using the structure of *Pc*Cel6A WT (PDB ID 5XCY) [26] as a model, and the structure was refined with PHENIX and Coot (v.0.9.8.91) [27]. Data collection and refinement statistics are given in Table 1. The root-mean-square deviations of the bond lengths of the Cα atoms between the structure of WT and mutants and the mean *B*-factors of atoms constituting the amino acids were calculated by the Structure Comparison program in PHENIX [28]. The dynamics of the enzymes were simulated with ensemble refinement in PHENIX [29]. The statistics of ensemble refinement are given in Table 3. The molecular graphics were prepared with the PyMOL Molecular Graphics System (v.2.5.5; Schrödinger, LLC), and a graph omitting part of the vertical axis was prepared with DeltaGraph (ver.7; Nihon Poladigital, K.K., Tokyo, Japan).

The following IDs contain the coordinates and associated structure factors of the catalytic domain of *Pc*Cel6A and its mutants deposited in the PDB database: 8WUP (WT), 8WW5 (C240S), 8WWT (C393S), and 8WX6 (C240S/C393S).

Phylogenic analysis. The amino acid sequences of characterized GH6 CBHs were collected with Batch Entrez (https://www.ncbi.nlm.nih.gov/sites/batchentrez) from the GenBank accession numbers listed in CAZy. The GH6 domain was extracted with DoMosaics [30] and aligned with MAFFT version 7 [31, 32]. Phylogeny was analyzed with MEGA X [33, 34] using the maximum likelihood method [35]. Unipro UGENE (v.48.1) [36] was used to check the number of target amino acids in the aligned sequence. The phylogenic tree was prepared with FigTree (v.1.4.4.; http://tree.bio.ed.ac.uk/software/figtree/). The graphics were prepared with PyMOL (v.2.5.5).

RESULTS AND DISCUSSION

Fluctuation of the substrate-enclosing loops in crystal structure.

The crystal structures of the catalytic domains of *Pc*Cel6A WT, C240S, C393S, and C240S/C393S were determined at 0.99, 1.01, 1.00, and 0.99 Å resolution, respectively (Table 1). All the crystals were classified into space group $P2_12_12_1$, and only one molecule was found in the asymmetric unit in agreement with the previously solved structure of *Pc*Cel6A WT (PDB ID 5XCY) [26]. The models were refined to R_{work} 14.7−15.9 % and R_{free} 15.4−16.9 %. The root-mean-square deviations of the bond lengths of the Cα atoms of mutants from the WT structure were within 0.4 Å (Table 2), and thus all of the structures can be regarded as quite similar [37, 38], indicating that the replacement of C240 and C393 with Ser did not distort the overall structure. The *B*-factors of the main chain (Fig. 1A) and side chain (Fig. 1B) normalized to the chain mean also showed a similar tendency among all four structures, and no marked difference was apparent at the 240th and 393rd amino acids.

Figure 2 illustrates the relative *B*-factors of WT and the mutants, with higher *B*-factors denoted by warmer colors and thicker ribbons. The *B*-factor of WT was relatively high at both the N-terminal loop (residue 172−177) and C-terminal loop (residue 390−425) (Fig. 2A). A similar trend was seen in C240S (Fig. 2B) and C393S (Fig. 2C), but C240S/C393S showed lower *B*-factor at both loops (Fig. 2D). These results indicated that removing both free Cys residues reduces the structural fluctuation of the substrate-enclosing loops.

Side chain variation in ensemble models.

Since the standard averaged structure masks information on minor conformations, we also performed ensemble refinement combined with molecular dynamics simulation. The possible states were divided into different models (Fig. 3A), and the values of R_{work} and R_{free} were lowered by 1.7−4.2 and 0.7−2.8 %, respectively (Table 3). Close-up

Table 1. X-ray data collection and refinement statistics for *Pc*Cel6A WT and mutants.

	WT	C240S	C393S	C240S/C393S
Data collection				
Beamline	PF BL-5A	PF BL-5A	PF BL-5A	PF BL-5A
Detector	Pilatus3 S6M	Pilatus3 S6M	Pilatus3 S6M	Pilatus3 S6M
Wavelength (Å)	1.0	1.0	1.0	1.0
Beam size (mm^2)	0.10 × 0.20	0.20 × 0.10	0.20 × 0.20	0.20 × 0.20
Positional change	0°–360°	0°–360°	0°–360°	0°–360°
Exposure time (s $frame^{-1}$)	0.1	0.1	0.1	0.1
Oscillation range (° $frame^{-1}$)	0.1	0.1	0.1	0.1
No. of total flames	3600	3600	3600	3600
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell parameters				
a (Å)	54.71	54.57	54.50	54.45
b (Å)	67.37	67.37	67.37	67.37
c (Å)	88.05	88.06	88.38	88.02
Resolution (Å)[a]	46.47–0.99 (1.01–0.99)	46.38–1.01 (1.03–1.01)	46.39–1.00 (1.02–1.00)	46.31–0.99 (1.01–0.99)
Total reflections	2,237,270	2,101,706	2,179,200	2,174,405
Unique reflections	180,424	169,019	175,525	179,971
Completeness (%)[a]	99.7 (97.6)	99.4 (97.0)	100.0 (99.5)	100.0 (99.7)
Redundancy (%)[a]	12.4 (10.9)	12.4 (11.9)	12.4 (11.5)	12.1 (10.7)
Average I/σ (I)[a]	25.6 (2.1)	26.1 (2.2)	33.1 (2.4)	22.5 (2.2)
R_{merge}[a]	0.052 (1.117)	0.055 (1.127)	0.038 (1.037)	0.049 (0.868)
Mosaicity (°)	0.127	0.156	0.106	0.175
Wilson *B*-factor ($Å^2$)	8.78	8.65	9.67	9.88
Refinement				
Resolution[a]	46.47–0.99 (1.03–0.99)	46.38–1.01 (1.05–1.01)	46.39–1.00 (1.04–1.00)	46.31–0.99 (1.03–0.99)
R_{work} (%)[a]	15.4 (23.0)	14.7 (22.1)	15.3 (22.3)	15.9 (22.1)
R_{free} (%)[a]	16.5 (25.3)	15.4 (22.7)	16.9 (22.8)	16.8 (25.1)
No. of reflections[a]	180,315 (17,557)	168,918 (16,357)	175,425 (17,280)	179,662 (17,746)
No. of atoms	4,323	4,292	4,552	3,597
Protein / Ligands	3,717 / 0	3,643 / 0	3,963 / 0	3,042 / 0
Solvent	606	649	589	555
RMSD from ideal values				
Bond length (Å)	0.005	0.005	0.005	0.005
Bond angles (°)	0.86	0.83	0.84	0.84
Ramachandran plot				
Favoured regions (%)	97.2	97.2	97.2	97.5
Additionally allowed (%)	2.5	2.8	2.8	2.5
Outliers (%)	0.3	0.0	0.0	0.0
PDB ID	8WUP	8WW5	8WWT	8WX6

Values in parentheses are for the highest resolution shell.

Table 2. Comparison of RMSD (Å) among the crystal structures of *Pc*Cel6A WT and mutants.

	WT	C240S	C393S	C240S/C393S
WT		0.351	0.312	0.366
C240S			0.378	0.404
C393S				0.416
C240S/C393S				

views around the 240th and 393rd amino acids (Figs. 3B−3E) show that the dihedral-angle distribution of the neighboring amino acids differed depending on the position of the free Cys substitution(s) (Fig. 4). Among the side chain of amino acids within 4 Å of the 240th residue, a prominent difference was seen in Q187 (Fig. 3B). The χ_1 angle of C240 was almost the same in WT and C393S (Figs. 4A and 4C), while that of S240 showed greater variation in C240S and C240S/C393S (Figs. 4B and 4D). This would influence the interaction with Q187 facing the 240th amino acid within hydrogen bond distance in the averaged models (Fig. 3C). The χ_3 angle of Q187 was almost within a 60° range in WT and C393S (Figs. 4A and 4C), whereas several kinds of minor conformations at Q187 were observed in C240S and C240S/C393S (Figs. 4B and 4D). Overall, the results suggest that steric hindrance between the two α-helixes within which S240 and Q187 are located might account for the lower activity of C240S at 60 °C than at 45 °C, as reported in the accompanying paper.

Substantial structural changes were not observed in the side chain of amino acids within 4 Å of the 393rd residue (Fig. 3D), but a possible interaction with the side chain of N362 was present. Although the distance between the side-chain sulfur or oxygen of the 393rd amino acid and the side-chain nitrogen of N362 was commonly about 3.4 Å in both WT and the mutants in the averaged models (Fig. 3E), the χ_1 and χ_2 angles in the ensemble models were different, depending on the nature of the 393rd amino acid. The distribution of these angles was relatively wide (a range of 40°) in C393S (Fig. 4C) and C240S/C393S (Fig. 4D). In comparison, a relatively larger proportion of models were distributed within 20° in WT (Fig. 4A) and C240S (Fig. 4B). Thus, the 393rd amino acid seems to interact with N362 in a similar

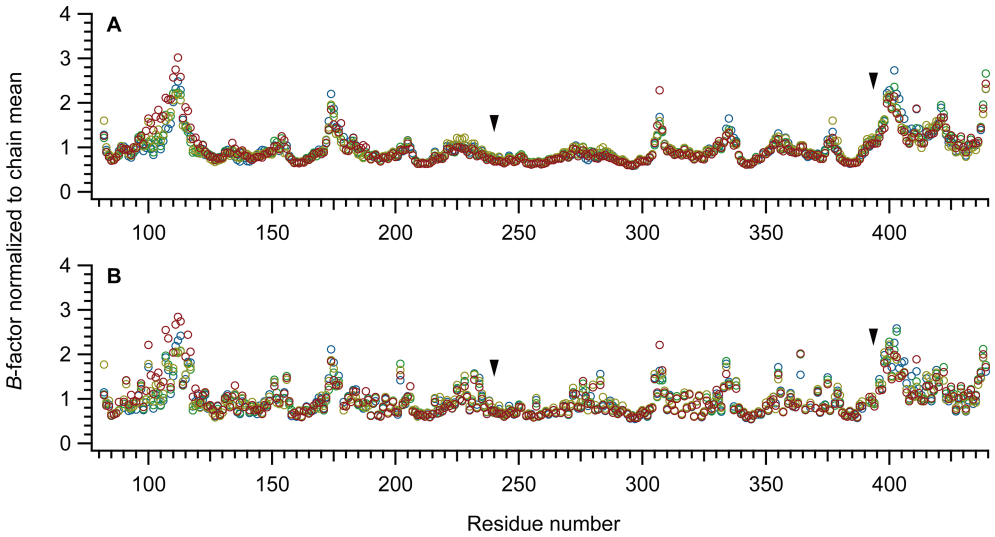

Fig. 1. *B*-factors of each residue's main chain (A) and side chain (B) normalized to the chain mean (WT, blue; C240S, green; C393S, yellow; C240S/C393S, red).
Arrowheads represent the positions of the 240th and 393rd amino acids.

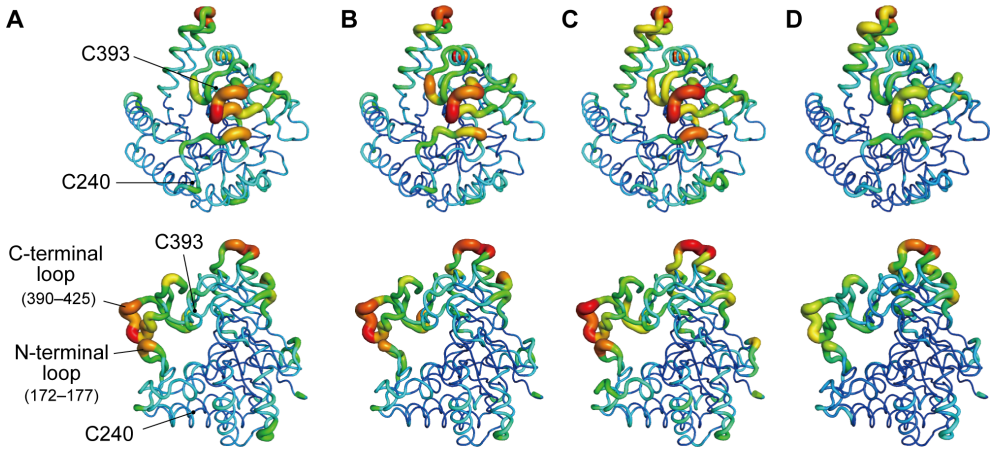

Fig. 2. *B*-factors in the crystal structures of *Pc*Cel6A (A) WT, (B) C240S, (C) C393S, and (D) C240S/C393S.
The warmer colors and bolder lines represent larger *B*-factors normalized in each structure. The figures below represent the views of the upper figures from the right side.

manner to the interaction of the 240th amino acid with Q187.

Stabilization of a disulfide bond on the C-terminal loop.

S393 in C240S/C393S is strictly directed towards the disulfide bond between C361 and C408, whereas a greater variety of conformations is seen in WT, C240S, and C393S (Fig. 3E). These results indicate that the conformation of the 393rd amino acid is also influenced by the 240th amino acid, possibly via some form of hydrogen bonding. The side-chain sulfur of C361, forming a disulfide bond with C408, lies within hydrogen-bonding distance from the side-chain nitrogen of N362 in the averaged structure (Fig. 3E). Although the χ_1 angle distribution of C361 and C408 did not show any significant difference among the four enzymes (Fig. 4), ensemble models of C240S/C393S appeared to be more parallel compared to WT (Fig. 3E). Possible reasons for this would include the smaller fluctuation of the C-terminal loop including C408 (Fig. 2) and the relatively fixed orientation of S393 (Fig. 3E).

As mentioned in the accompanying paper, deriving from the results of the thermal shift assay, the effect of C240 on a carboxyl-carboxylate hydrogen bond may also be involved in the phenomenon. There are four carboxylic acid pairs in *Pc*Cel6A WT (E101−E392, D359−D412, D165−E179, and D170−D216), and two of which (E101−E392, D359−D412) are at the root of the C-terminal loop (Fig. 5A). E392, paired with E101, is next to C393, and D359−D412 is relatively close to C361−C408 disulfide bond as demonstrated by the distances of the Cα atoms between D359 and C361, D359 and C408, D412 and C361, and D412 and C408, all being less than 9 Å in WT. Comparing the left-side figures (WT and C393S) with the right-side figures (C240S and C240S/C393S) in the close-up of the ensemble models of the E101−E392 pair (Fig. 5B) and the D359−D412 pair (Fig. 5C), the proportion of side chains pointing with each other appeared to slightly increase by replacing C240 with Ser. On the other hand, such a tendency was not seen in D165−E179 (Fig. 5D) and D170−D216 (Fig. 5E), despite their relative proximity to C240 within the molecule, as indicated by the distances between the Cα atoms of D165 and C240 being less than 10 Å, and between D216 and C240 being less than 14 Å in WT. Therefore, even though it is still challenging to determine

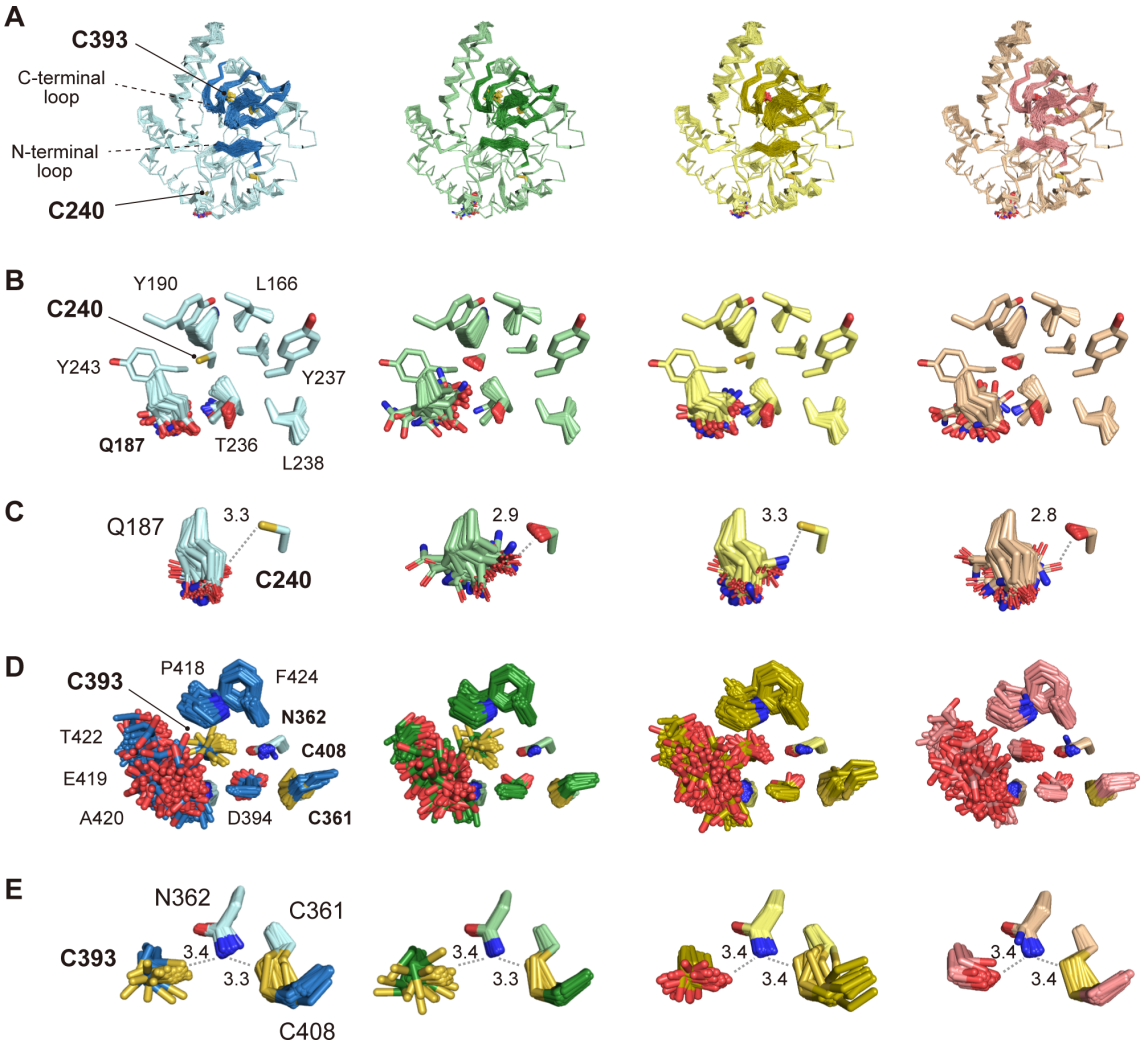

Fig. 3. Ensemble models of the crystal structures of *Pc*Cel6A WT (blue), C240S (green), C393S (yellow), and C240S/C393S (pink). The N- and C-terminal loops are shown in darker color.

(A) The side chains of Cys residues forming disulfide bonds, the 240th and 393rd amino acids, and Q187 are shown as sticks in the overall structure. (B) Residues within 4 Å from the 240th amino acid. (C) Extracted view of 240th amino acid and the adjacent Q187. Distance (Å) between the sidechains in a static structure is indicated by a dotted line and a single decimal number. (D) Residues within 4 Å from the 393rd amino acid and the C361−C408 disulfide bond. (E) Extracted view of a pathway from the 393rd amino acid to the C361−C408 disulfide bond via N362. Distance (Å) between the sidechains in a static structure is indicated by dotted lines and single decimal numbers.

Table 3. Statistics of the ensemble refinement for the crystal structures of *Pc*Cel6A WT and mutants.

	WT	C240S	C393S	C240S/C393S
Refinement parameters				
Relaxation time (ps)	2.0	2.0	2.0	2.0
pTLS (%)	0.6	0.6	0.6	0.6
Conformers (#)	200	100	200	200
Refinement and model statistics				
Resolution range (Å)	46.47–0.99 (1.01–0.99)	46.38–1.01 (1.03–1.01)	46.39–1.00 (1.02–1.00)	46.31–0.99 (1.01–0.99)
R_{work} (%)	12.9 (21.2)	13.0 (20.9)	13.0 (20.6)	12.6 (19.2)
R_{free} (%)	14.0 (23.5)	14.7 (25.7)	15.1 (23.4)	14.0 (18.8)
ΔR_{work} (%)	– 2.5	– 1.7	– 2.3	–4.2
ΔR_{free} (%)	– 2.5	– 0.7	– 1.8	– 2.8
Mean RMSD per structure				
Bonds (Å)	0.008	0.008	0.008	0.008
Angles (°)	1.110	1.099	1.104	1.108
Dihedrals (°)	16.27	16.31	16.15	16.18

Values in parentheses are for the highest resolution shell. pTLS is the proportion of atoms included in the process of TLS fitting. ΔR_{work} and ΔR_{free} represent the value of the ensemble structure minus the single structure.

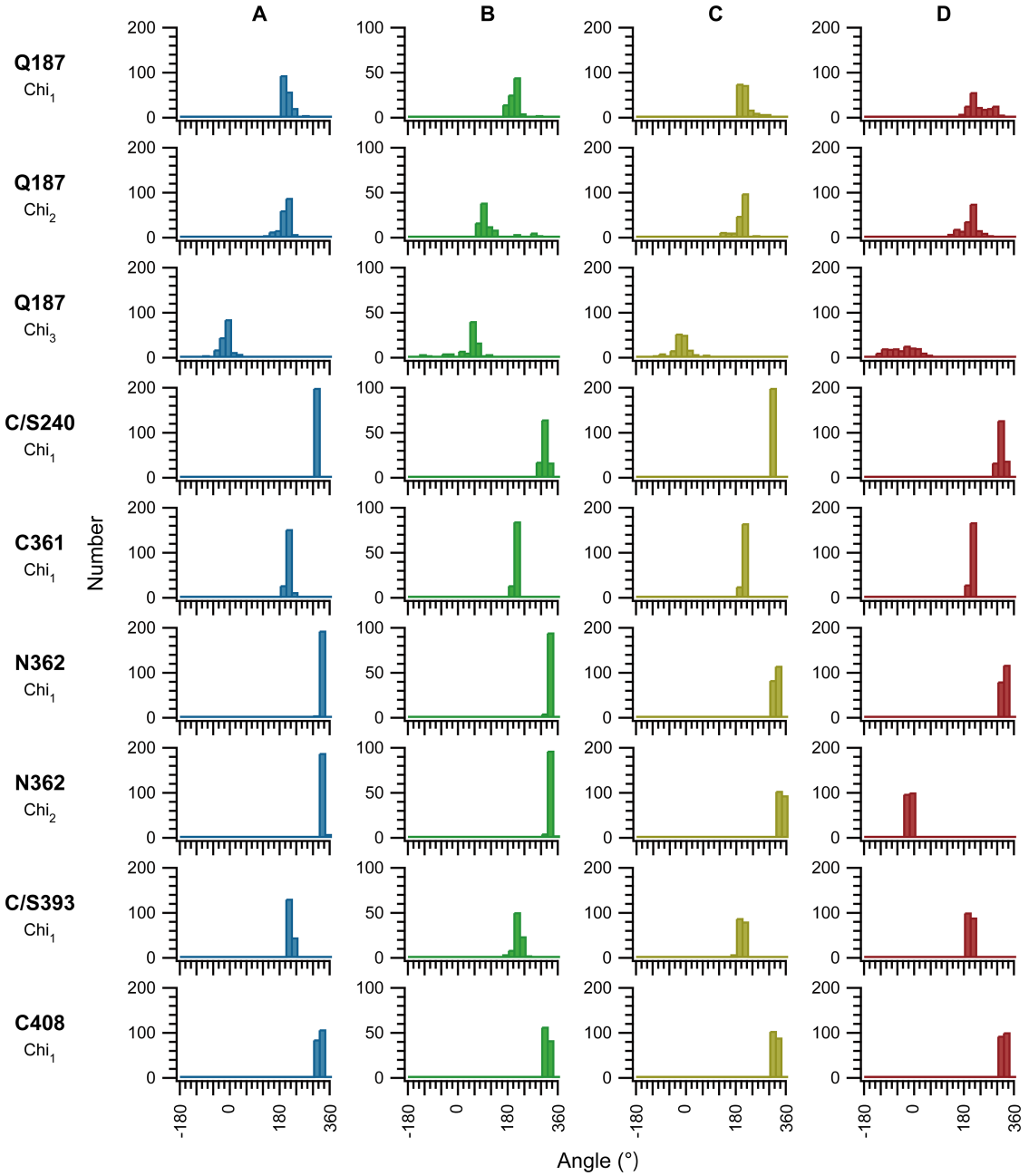

Fig. 4. Histograms of the dihedral angles of the 240th and 393rd amino acids and some neighboring residues in the ensemble models of WT (A), C240S (B), C393S (C), and C240S/C393S (D).

the detailed pathway, it seems possible that the C240 substitution contributed to maintaining the E101−E392 and D359−D412 pairs, leading to stabilizing the network from the 393rd amino acid to C361−C408 (Fig. 4E).

Moreover, the intermolecular interaction might also occur between C240 and the carboxylic acid pairs, as discussed in the accompanying paper. Among the four carboxylic acid pairs in *Pc*Cel6A WT, three, except for the pair of D170 and the acid catalyst D216, are located near the outside of the enzyme (Fig. 5A). Notably, the side chains of E101 in Fig. 5B and D412 in Fig. 5C, which are located on the enzyme surface, exhibited a wide range of models. Thus, there may be the possibility that these carboxylic acids and C240, positioned near the enzyme surface in two independent molecules, react intermolecularly at high temperatures. Conversely, substituting C240 possibly kept the strong carboxyl-carboxylate hydrogen bond, bringing the higher T_m at pH 3−6 to C393S and C240S/C393S (Figs. 5C and 5D in the accompanying paper). In any event, hydrogen bonding in the loop region is expected to enhance thermotolerance [39].

Strength of hydrogen bonds in the C-terminal loop.

Focusing on the environment around the 393rd amino acid, located in the C-terminal loop, the distance from the side-chain sulfur or oxygen of the 393rd amino acid to the main chains of neighboring residues (D394−N, G395−N, G395−O, and P418−O) was 0.3 or 0.4 Å shorter in C393S and C240S/C393S than in WT and C240S, indicating stronger hydrogen bonding (Fig. 6). C393S and C240S/C393S both displayed higher activity at 60 ˚C than at 45 ˚C, as described in the accompanying paper (Fig. 3 in the accompanying paper), supporting the idea that the strength of hydrogen bonds involving the main chains of amino acids

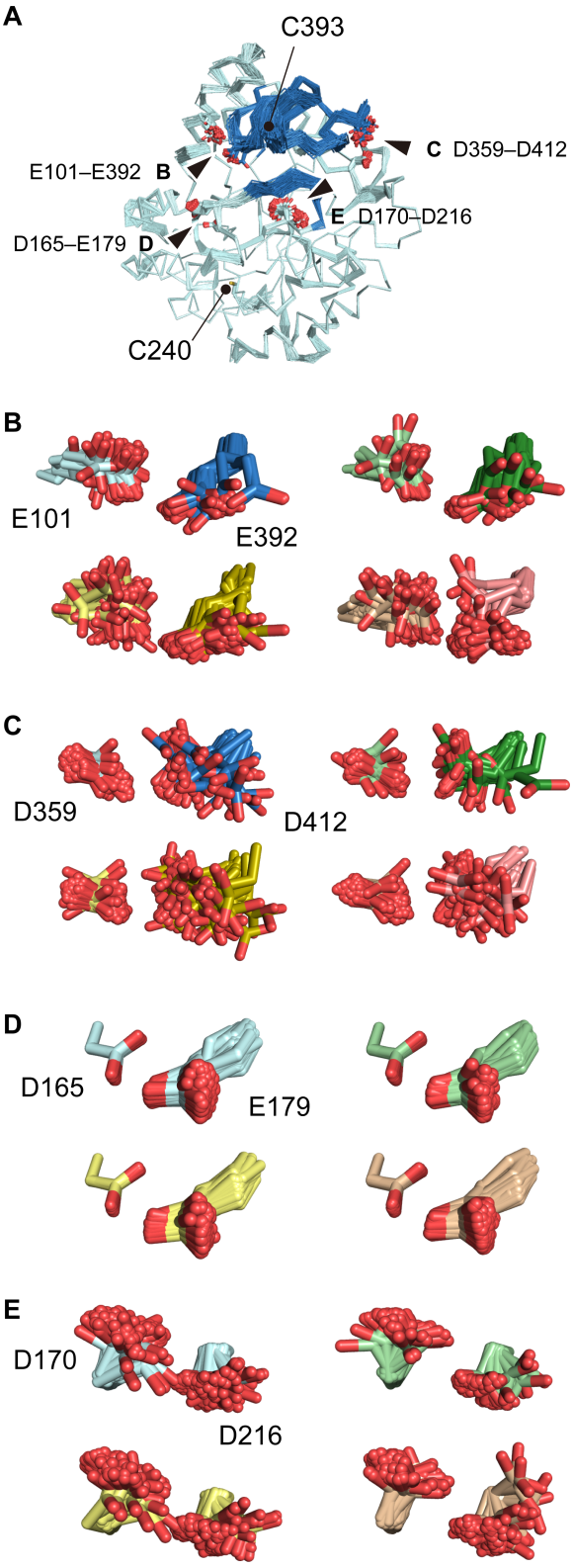

Fig. 5. Carboxylic acid pairs in *Pc*Cel6A.

(A) Potential carboxyl-carboxylate hydrogen bond sites. Ensemble models of the pairs and free Cys residues are shown as sticks. The N- and C-terminal loops are colored darker. The bold one-letter alphabet indicates the name of the enlarged figure seen from the arrowhead. (B–E) Magnified view of each carboxylic acid pair in WT (blue), C240S (green), C393S (yellow), and C240S/C393S (pink).

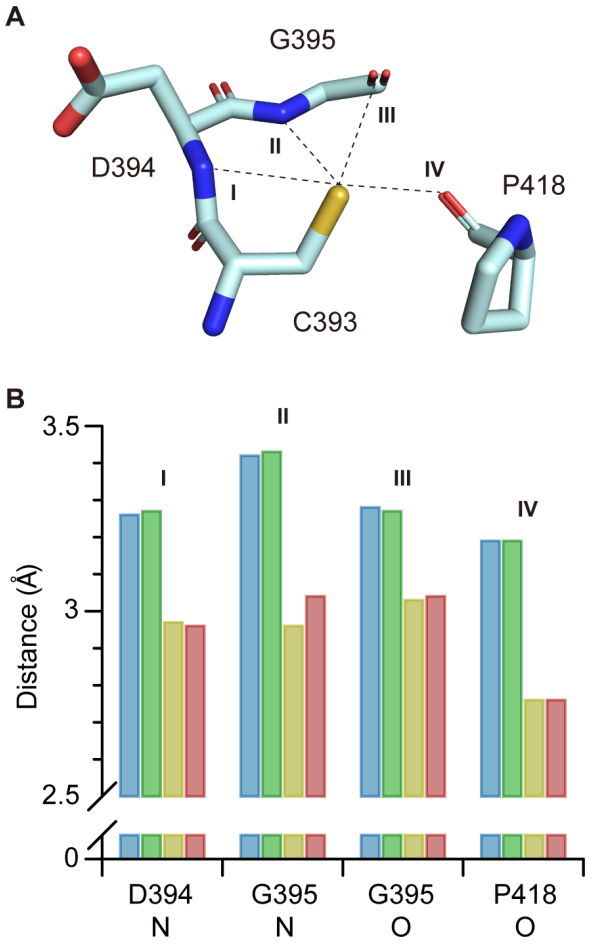

Fig. 6. Interaction of the C393 side chain with the main chain of neighboring residues.

(A) The close-up view around C393 in WT. The dashed line is the measured length from the side chain of the 393rd amino acid to the main-chain nitrogen of D394 (I) and G395 (II), and the main-chain oxygen of G395 (III) and P418 (IV). (B) Distance between the side chain of the 393rd amino acid of WT (blue), C240S (green), C393S (yellow), and C240S/C393S (red) with the main chains of neighboring residues illustrated in the structure.

in the C-terminal loop is a critical determinant of the thermotolerance of *Pc*Cel6A.

A contribution of stronger hydrogen bonding to thermotolerance was previously predicted based on modeling of GH6 CBH from the ascomycete *H. insolens* (*Hi*Cel6A), whose activity at high temperature is increased by a substitution corresponding to *Pc*Cel6A C393S [14]. In that study, hydrogen atoms and the substitution corresponding to *Pc*Cel6A C393S were introduced into the X-ray crystal structure of *Hi*Cel6A [40] by optimizing the hydrogen-bond network and side-chain packing with REDUCE [41] and SHARPEN [42], respectively. These findings are consistent with the idea that stronger hydrogen bonds in the C-terminal loop increase the thermotolerance of other GH6 CBHs, not just *Pc*Cel6A.

Conservation of free Cys in fungal GH6 CBH.

To see if GH6 CBHs from other organisms also possess free Cys residues, we examined amino acids corresponding to C240 and C393 in *Pc*Cel6A on the phylogenetic tree of characterized GH6 members (Fig. 7). Interestingly, free Cys at positions corresponding to the 240th and 393rd amino acids of *Pc*Cel6A were only conserved among aerobic fungal CBH, but not among anaerobic fungal CBHs, fungal EGs, bacterial CBHs, or bacterial EGs. Some aerobic fungal GH6 CBHs contain different amino acids at the positions corresponding to the free Cys of *Pc*Cel6A. Examples include

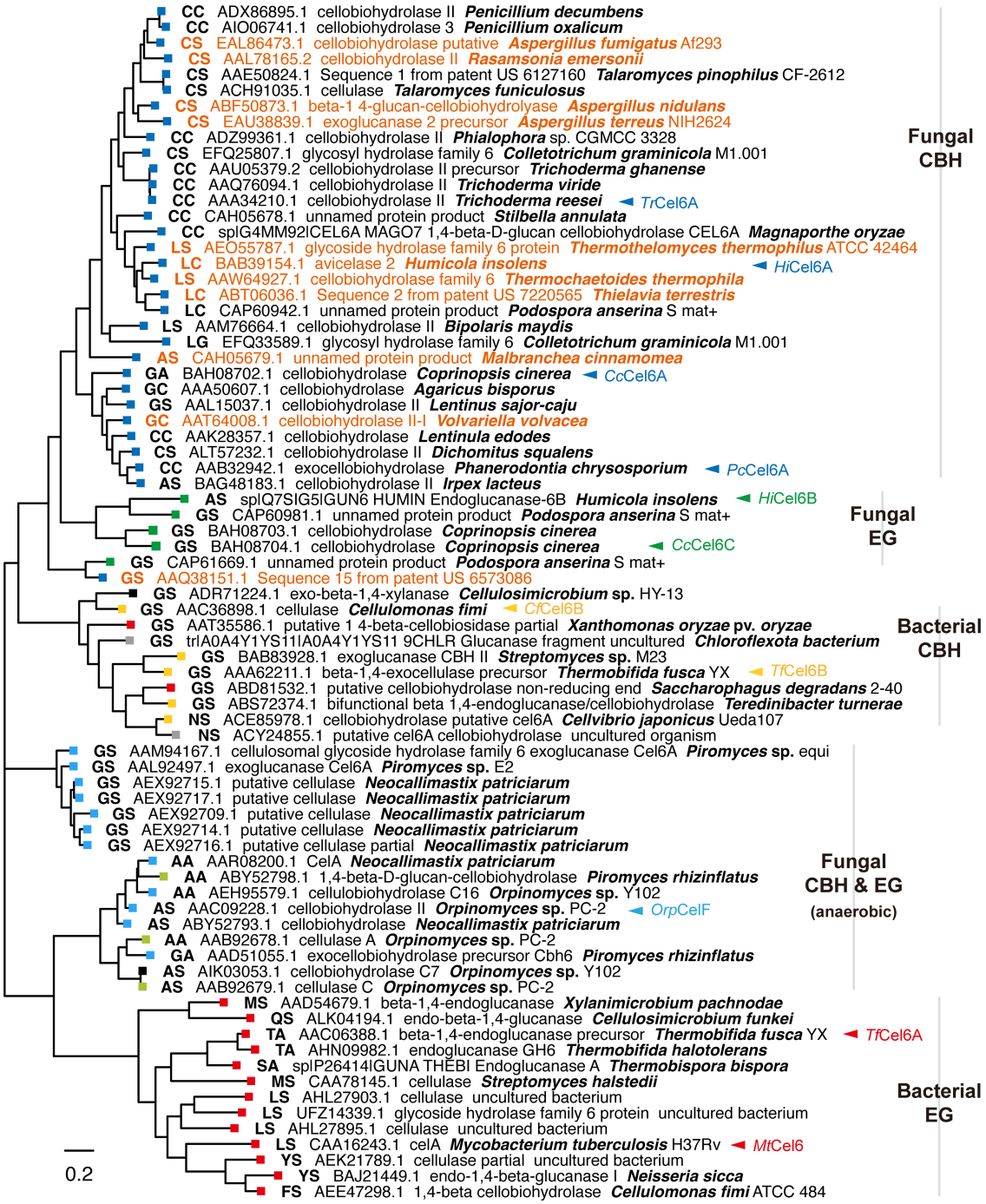

Fig. 7. Conservation of free Cys among characterized GH6.

The colors of the node tips indicate fungal GH6 described as EC 3.2.1.91 in CAZy (blue), eukaryotic GH6 for which EC 3.2.1.4 activity was reported (green), and bacterial GH6 described as EC 3.2.1.91 in CAZy (orange), GH6 from an anaerobic fungus characterized as EC 3.2.1.91 in CAZy (light blue), GH6 from an anaerobic fungus for which EC 3.2.1.4 was reported (light green), bacterial GH6 described as EC 3.2.1.4 in CAZy (red), GH6 from an unclassified organism (gray), and GH6 indicated as other than EC 3.2.1.4 and 3.2.1.91 in CAZy (black). The first and second letters before the file name represent the amino acids at the positions corresponding to the 240th and 393rd free Cys residues in *Pc*Cel6A. The label of GH6 CBH from an aerobic thermophilic fungus is written in dark orange. The abbreviated names of some well-characterized GH6s are shown after the arrowheads: *Tr*Cel6A (*Trichoderma reesei* CBH), *Hi*Cel6A *Humicola insolens* CBH), *Cc*Cel6A (*Coprinopsis cinerea* CBH), *Hi*Cel6A (*H. insolens* EG), *Cc*Cel6C (one of *C. cinerea* EG), *Cf*Cel6B (*Cellulomonas fimi* CBH), *Tf*Cel6B (*Thermobifida fusca* CBH), *Orp*CelF (one of the *Orpinomyces* sp. CBHs), *Tf*Cel6A (*T. fusca* EG), and *Mt*Cel6 (*Mycobacterium tuberculosis* EG).

thermophilic fungi, such as *Aspergillus fumigatus* [43], *Rasamsonia emersonii* [44], *A. nidulans* [45], *A. terreus* [46], *Thermothelomyces thermophilus* [47], *H. insolens* [48], *Thermochaetoides thermophila* [49], *Thielavia terrestris* [49], *Malbranchea cinnamomea* [49], *Volvariella volvacea* [50], and *Chrysosporium lucknowense* [51]. Although the conservation of GH6s' free Cys corresponding to C240 and C393 in *Pc*Cel6A appeared different according to the oxygen abundance and temperature of the environment where the host organisms live, the causal relationships have not been clarified and other factors may also be involved.

CONCLUSION

The mechanistic basis of the thermotolerance of the double mutant, C240S/C393S, of GH6 CBH from the basidiomycete *Phanerochaete chrysosporium* (*Pc*Cel6A), which lacks the free Cys residues, was investigated by comparing the X-ray crystal structures of WT, C240S/C393S and the individual mutants, C240S and C393S. Our results suggest that the increased activity of C393S and C240S/C393S at an elevated temperature compared to the WT enzyme is predominantly due to the stabilization of the substrate-enclosing loops. The C393S substitution also plays a greater role than the C240S substitution in suppressing thermal denaturation. The overall structures of WT and all three mutants were very similar. However, C240S/C393S had the lowest *B*-factor at the N-terminal and C-terminal substrate-enclosing loops, suggesting that reduced structural fluctuation at these locations, possibly due at least in part to stronger hydrogen bonding involving the 393rd residue, could be the main contributor to the increased thermotolerance.

CONFLICTS OF INTERESTS

The authors declare that they have no competing interests.

ACKNOWLEDGMENTS

We thank all the beamline staff members at the High Energy Accelerator Research Organization and Dr. Takuya Ishida at IMEQRD Co., Ltd. for their help in collecting the X-ray crystal structure data. This work was performed with the approval of the Photon Factory Program Advisory Committee (Proposal No. 2021G521). SY is grateful for financial support from the UTokyo Sustainable Agriculture Education Program during a master's course. This study received financial support from Grants-in-Aid for Scientific Research (A) from the Japan Society for the Promotion of Science (JSPS) (No. 23H00341 to KI), the Sasakawa Scientific Research Grant from The Japan Science Society (No. 2021-4060 to SY), and a Grant-in-Aid for JSPS Fellows (No. 22J12651 to SY).

REFERENCES

[1] Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014; 42: 490−5.

[2] Boisset C, Fraschini C, Schülein M, Henrissat B, Chanzy H. Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase Cel6A from *Humicola insolens* and its mode of synergy with cellobiohydrolase Cel7A. Appl Environ Microbiol. 2000; 66: 1444−52.

[3] Abuja PM, Pilz I, Claessens M, Tomme P. Domain structure of cellobiohydrolase II as studied by small angle X-ray scattering: close resemblance to cellobiohydrolase I. Biochem Biophys Res Commun. 1988; 156: 180−5.

[4] Nakamura A, Tasaki T, Ishiwata D, Yamamoto M, Okuni Y, Visootsat A, et al. Single-molecule imaging analysis of binding, processive movement, and dissociation of cellobiohydrolase *Trichoderma reesei* Cel6A and its domains on crystalline cellulose. J Biol Chem. 2016; 291: 22404−13.

[5] Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, et al. Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. Science. 2011; 333: 1279−82.

[6] Medve J, Ståhlberg J, Tjerneld F. Adsorption and synergism of cellobiohydrolase I and II of *Trichoderma reesei* during hydrolysis of microcrystalline cellulose. Biotechnol Bioeng. 1994; 44: 1064−73.

[7] Badino SF, Christensen SJ, Kari J, Windahl MS, Hvidt S, Borch K, et al. Exo-exo synergy between Cel6A and Cel7A from *Hypocrea jecorina*: Role of carbohydrate binding module and the endo-lytic character of the enzymes. Biotechnol Bioeng. 2017; 114: 1639−47.

[8] Vršanská M, Biely P. The cellobiohydrolase I from *Trichoderma reesei* QM 9414: action on cello-oligosaccharides. Carbohydr Res. 1992; 227: 19−27.

[9] Voutilainen SP, Puranen T, Siika-Aho M, Lappalainen A, Alapuranen M, Kallio J, et al. Cloning, expression, and characterization of novel thermostable family 7 cellobiohydrolases. Biotechnol Bioeng. 2008; 101: 515−28.

[10] Wang XJ, Peng YJ, Zhang LQ, Li AN, Li DC. Directed evolution and structural prediction of cellobiohydrolase II from the thermophilic fungus *Chaetomium thermophilum*. Appl Microbiol Biotechnol. 2012; 95: 1469−78.

[11] Heinzelman P, Snow CD, Wu I, Nguyen C, Villalobos A, Govindarajan S, et al. A family of thermostable fungal cellulases created by structure-guided recombination. Proc Natl Acad Sci USA. 2009; 106: 5610−5.

[12] Wu I, Arnold FH. Engineered thermostable fungal Cel6A and Cel7A cellobiohydrolases hydrolyze cellulose efficiently at elevated temperatures. Biotechnol Bioeng. 2013; 110: 1874−83.

[13] Ito Y, Ikeuchi A, Imamura C. Advanced evolutionary molecular engineering to produce thermostable cellulase by using a small but efficient library. Protein Eng Des Sel. 2013; 26: 73−9.

[14] Heinzelman P, Snow CD, Smith MA, Yu X, Kannan A, Boulware K, et al. SCHEMA recombination of a fungal cellulase uncovers a single mutation that contributes markedly to stability. J Biol Chem. 2009; 284: 26229−33.

[15] Wu I, Heel T, Arnold FH. Role of cysteine residues in thermal inactivation of fungal Cel6A cellobiohydrolases. BBA - Proteins Proteom. 2013; 1834: 1539−44.

[16] Yamaguchi S, Sunagawa N, Tachioka M, Igarashi K, Samejima M. Thermostable mutants of glycoside hydrolase family 6 cellobiohydrolase from the basidiomycete *Phanerochaete chrysosporium*. J Appl Glycosci. 2020; 67: 79−86.

[17] Yamaguchi S, Sunagawa N, Samejima M, Igarashi K. Thermotolerance Mechanism of Fungal GH6 Cellobiohydrolase. Part I. Characterization of Thermotolerant Mutant from the Basidiomycete *Phanerochaete chrysosporium*. J Appl Glycosci. 2024; 71: 55−62.

[18] Yamaguchi S, Sunagawa N, Matsuyama K, Tachioka M, Hirota E, Takahashi S, et al. Preparation of large-volume crystal of cellulase under microgravity to investigate the mechanism of thermal stabilization. Int J Microgravity Sci Appl. 2021; 38: 380103.

[19] Nakamura H, Fukuyama S, Yoshizaki I, Yoda SI. Theory of vapor diffusion in the high-density protein crystal growth

device and its application to the measurement of the vapor pressures of aqueous solutions. J Cryst Growth. 2003; 259: 149−59.
[20] Benvenuti M, Mangani S. Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. Nat Protoc. 2007; 2: 1633−51.
[21] Kabsch W. XDS. Acta Crystallogr Sect D Biol Crystallogr. 2010; 66: 125−32.
[22] Evans P. Scaling and assessment of data quality. Acta Crystallogr Sect D Biol Crystallogr. 2006; 62: 72−82.
[23] Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr Sect D Biol Crystallogr. 2013; 69: 1204−14.
[24] Yamada Y, Matsugaki N, Chavas LMG, Hiraki M, Igarashi N, Wakatsuki S. Data management system at the photon factory macromolecular crystallography beamline. J Phys Conf Ser. 2013; 425: 012017.
[25] Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. Acta Crystallogr Sect D Struct Biol. 2019; 75: 861−77.
[26] Tachioka M, Nakamura A, Ishida T, Igarashi K, Samejima M. Crystal structure of a family 6 cellobiohydrolase from the basidiomycete *Phanerochaete chrysosporium*. Acta Crystallogr Sect F Struct Biol Commun. 2017; 73, 398−403.
[27] Emsley P, Lohkamp B. Features and development of *Coot.* Acta Crystallogr Sect D Biol Crystallogr. 2010; 66: 486−501.
[28] Moriarty NW, Liebschner D, Klei HE, Echols N, Afonine PV, Headd JJ, et al. Interactive comparison and remediation of collections of macromolecular structures. Protein Sci. 2018; 27: 182−94.
[29] Burnley BT, Afonine PV, Adams PD, Gros P. Modelling dynamics in protein crystal structures by ensemble refinement. Elife. 2012; 1: e00311.
[30] Moore AD, Heldy A, Terrapon N, Weiner J, Bornberg-Bauer E. DoMosaics: Software for domain arrangement visualization and domain-centric analysis of proteins. Bioinformatics. 2014; 30: 282−3.
[31] Kuraku S, Zmasek CM, Nishimura O, Katoh K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic Acids Res. 2013; 41: 22−8.
[32] Katoh K, Rozewicki J, Yamada KD. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019; 20: 1160−6.
[33] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018; 35: 1547−9.
[34] Stecher G, Tamura K, Kumar S. Molecular evolutionary genetics analysis (MEGA) for macOS. Mol Biol Evol. 2020; 37: 1237−9.
[35] Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Bioinformatics. 1992; 8: 275−82.
[36] Okonechnikov K, Golosova O, Fursov M, Varlamov A, Vaskin Y, Efremov I, et al. Unipro UGENE: A unified bioinformatics toolkit. Bioinformatics. 2012; 28: 1166−7.
[37] Stark A, Sunyaev S, Russell RB. A model for statistical significance of local similarities in structure. J Mol Biol. 2003; 326: 1307−16.
[38] Bin Zhou R, Lu HM, Liu J, Shi JY, Zhu J, Lu QQ, et al. A systematic analysis of the structures of heterologously expressed proteins and those from their native hosts in the RCSB PDB archive. PLoS One, 2016; 11: e0161254.
[39] Tang H, Shi K, Shi C, Aihara H, Zhang J, Du G. Enhancing subtilisin thermostability through a modified normalized B-factor analysis and loop-grafting strategy. J Biol Chem. 2019; 294: 18398−407.
[40] Varrot A, Macdonald J, Stick RV, Pell G, Gilbert HJ, Davies GJ. Distortion of a cellobio-derived isofagomine highlights the potential conformational itinerary of inverting β-glucosidases. Chem Commun. 2003; 3: 946−7.
[41] Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol. 1999; 285: 1735−47.
[42] Loksha IV, Maiolo II JRI, Hong CW, Ng A, Snow CD. SHARPEN—Systematic hierarchical algorithms for rotamers and proteins on an extended network. J Comput Chem. 2009; 30: 999−1005.
[43] Couger B, Weirick T, Damásio ARL, Segato F, Polizeli MDLTDM, de Almeida RSC, et al. The genome of a thermo tolerant, pathogenic albino *Aspergillus fumigatus*. Front Microbiol. 2018; 9: 1−11.
[44] Houbraken J, Spierenburg H, Frisvad JC. *Rasamsonia*, a new genus comprising thermotolerant and thermophilic *Talaromyces* and *Geosmithia* species. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol. 2012; 101: 403−21.
[45] Lu H, Luo H, Shi P, Huang H, Meng K, Yang P, Yao B. A novel thermophilic endo-β-1,4-mannanase from *Aspergillus nidulans* XZ3: Functional roles of carbohydrate-binding module and Thr/Ser-rich linker region. Appl Microbiol Biotechnol. 2014; 98: 2155−63.
[46] Gao J, Weng H, Zhu D, Yuan M, Guan F, Xi Y. Production and characterization of cellulolytic enzymes from the thermoacidophilic fungal *Aspergillus terreus* M11 under solid-state cultivation of corn stover. Bioresour Technol. 2008; 99: 7623−9.
[47] Roy SK, Dey SK, Raha SK, Chakrabarty SL. Purification and properties of an extracellular endoglucanase from *Myceliophthora thermophila* D-14 (ATCC 48104). J Gen Microbiol. 1990; 136: 1967−71.
[48] Hayashida S, Yoshioka H. Production and purification of thermostable cellulases from *Humicola insolens* YH-8. Agric Biol Chem. 1980; 44: 1721−28.
[49] Maheshwari R, Bharadwaj G, Bhat MK. Thermophilic fungi: their physiology and enzymes. Microbiol Mol Biol Rev. 2000; 64: 461−88.
[50] Zhao X, Yu C, Zhao Y, Liu S, Wang H, Wang C, Guo L, Chen M. Changes in mannitol content, regulation of genes involved in mannitol metabolism, and the protective effect of mannitol on *Volvariella volvacea* at low temperature. Biomed Res Int. 2019; 2019: 1−12.
[51] Singh B. *Myceliophthora thermophila* syn. *Sporotrichum thermophile*: A thermophilic mould of biotechnological potential. Crit Rev Biotechnol. 2016; 36: 59−69.