J. Appl. Glycosci., **71**, 37−46 (2024)
doi:10.5458/jag.jag.JAG-2023_0011

Regular Paper

Effect of Free Cysteine Residues to Serine Mutation on Cellodextrin Phosphorylase

(Received October 3, 2023; Accepted January 19, 2024)

Tomohiro Kuga,[1] Naoki Sunagawa,[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: Cellodextrin phosphorylase (CDP) plays a key role in energy-efficient cellulose metabolism of anaerobic bacteria by catalyzing phosphorolysis of cellodextrin to produce cellobiose and glucose 1-phosphate, which can be utilized for glycolysis without consumption of additional ATP. As the enzymatic phosphorolysis reaction is reversible, CDP is also employed to produce cellulosic materials *in vitro*. However, the enzyme is rapidly inactivated by oxidation, which hinders *in vitro* utilization in aerobic environments. It has been suggested that the cysteine residues of CDP, which do not form disulfide bonds, are responsible for the loss of activity, and the aim of the present work was to test this idea. For this purpose, we replaced all 11 free cysteine residues of CDP from *Acetivibrio thermocellus* (formerly known as *Clostridium thermocellum*) with serine, which structurally resembles cysteine in our previous work. Herein, we show that the resulting CDP variant, named CDP-CS, has comparable activity to the wild-type enzyme, but shows increased stability to oxidation during long-term storage. X-Ray crystallography indicated that the mutations did not markedly alter the overall structure of the enzyme. Ensemble refinement of the crystal structures of CDP and CDP-CS indicated that the C372S and C625S mutations reduce structural fluctuations in the protein main chain, which may contribute to the increased stability of CDP-CS to oxidation.

Key words: Cellulose conversion, cellodextrin phosphorylase, stability, cysteine, oxidation

INTRODUCTION

Cellodextrin phosphorylase (CDP, EC 2.4.1.49), a member of glycoside hydrolase family 94 (GH94) in the CAZy (Carbohydrate Active Enzymes) database [1], is a key enzyme for cellulose metabolism by anaerobic bacteria. CDP catalyzes the phosphorolysis of soluble cellodextrin by utilizing phosphate ion, which exists abundantly in the cell, to afford cellobiose and several α-glucose 1-phosphate (αG1P) molecules, depending on the degree of polymerization (DP) of the substrate [2, 3]. This cellulose metabolic pathway is energetically efficient, because αG1P can be converted to glucose 6-phosphate (G6P) by phosphoglucomutase in the glycolytic pathway without consumption of ATP. In contrast, the hydrolytic cleavage of cellulose by β-glucosidase yields glucose molecules, which can be converted to G6P by hexokinase via a process that requires one mole of ATP per mole of glucose. The energy-efficiency of CDP-mediated cellulose metabolism is advantageous for anaerobic bacteria, as resources and energy are limited in the anaerobic environment [4, 5].

CDP can also catalyzes the reverse phosphorolysis reaction (synthetic reaction) *in vitro* [6]. In this reaction, one glucose residue from αG1P is non-processively transferred to the non-reducing end of cellodextrin, and the product becomes the next substrate. This iterative synthetic reaction can generate insoluble cellodextrin from cellobiose or glucose as a glycosyl acceptor and αG1P as a glycosyl donor [7]. Cellulose thus obtained from cellobiose had cellulose II crystalline form with a degree of polymerization (DP) < 10 [7]. In other studies, cellulose synthesized by CDP was reported to form lamellar platelet crystals with high crystallinity [8] and to exhibit a network-like higher-order structure [9−11]. Furthermore, CDP can accommodate a variety of acceptor and donor substrates, affording products with potential applications as functional oligosaccharides or cellulosic materials; examples include β-(1→4) hetero-D-glucose and D-xylose-based oligosaccharides and C1-modified cellulose [12−14].

Thus, CDP has a variety of potential applications for energy-efficient cellulose degradation and the production of green materials. However, the stability of the protein is insufficient for industrial use. For example, we encountered poor stability of CDP during our attempts to synthesize cellulose on the International Space Station for one month [10], and purified CDP expressed in *Escherichia coli* lost its activity during reaction or storage for several weeks. Thus, the aim of the present work was to develop a strategy to improve the stability of CDP. In this context, it has been

[†]Corresponding author (Tel. +81−3−5841−5255, E-mail: aquarius@mail.ecc.u-tokyo.ac.jp, ORCID ID: https://orcid.org/0000-0001-5152-7177)

Abbreviations: αG1P, α-glucose 1-phosphate; CDP, cellodextrin phosphorylase; DLS, dynamic light scattering; DP, degree of polymerization; G6P, glucose 6-phosphate; RMSD, root-mean-square deviation; RMSF, root-mean-square fluctuation.

reported that the activity of CDP was restored by adding sulfur-containing reducing agents such as dithiothreitol to the reaction mixture [6, 15]. We therefore considered that the oxidation of free cysteine residues, not involved in disulfide bonds, might impact on CDP's activity. Free cysteines can negatively affect the stability of proteins; for example, free cysteines decrease the thermal stability and induce aggregation of antibodies [16, 17]. In addition, substitution of free cysteine residues improved the stability of cellobiohydrolase [18, 19], feruloyl esterase [20], carboxyesterase [21] and formate dehydrogenase [22]. The recently solved crystal structure of CDP from *Acetivibrio thermocellus,* formerly known as *Clostridium thermocellum,* revealed the presence of 11 free cysteines [23], which might explain the poor stability of CDP.

In our previous paper, we created a CDP variant whose all 11 free cysteine residues of CDP from *A. thermocellus* with serine, in order to stabilize CDP against oxidation for a 1-month reaction on the International Space Station [10]. There, we used this CDP variant, designated as CDP-CS, to synthesize cellulose and discussed how gravity influenced higher-ordered structure of *in vitro* synthesized cellulose. Hereafter, CDP was referred to as CDP-WT to distinguish it from CDP-CS in this paper. The aim of this paper, herein, was to describe the effect of 11 mutations on the stability to oxidation and thermal stability of CDP-CS. Furthermore, we determined the structure of CDP-CS by X-ray crystallography and discussed the effect of 11 cysteine-to-serine mutations from the structural aspect. Ensemble refinement of the crystal structures of CDP-WT and CDP-CS indicated that the C80S mutation was associated with increased structural fluctuation, while the C372S and C625S mutations were associated with decreased structural fluctuation. These findings should be helpful in developing a rational strategy to increase the oxidation resistance of CDP-WT and related enzymes.

MATERIALS AND METHODS

Materials. BugBuster and Overnight Express medium were purchased from Merck KGaA (Darmstadt, Germany). D-Cellobiose and αG1P disodium salt were purchased from Sigma-Aldrich Co. LLC (St. Louis MO, USA). D-Cellotriose was purchased from Tokyo Chemical Industry Corporation (Tokyo, Japan). Other chemicals were purchased from FUJIFILM Wako Pure Chemicals Corporation (Osaka, Japan). *E. coli* BL21 (DE3) competent cells were purchased from Nippon Gene (Tokyo, Japan).

Enzyme preparation. A pET-28a vector harboring *cdp-wt* gene from *A. thermocellus* YM4 was kindly provided by Professor Motomitsu Kitaoka of Niigata University, prepared as described previously [24]. There, N terminal residues of CDP-WT were changed from MIT (atgattact) to MVT (atggtact), probably because of the introduction of *Nco*I restriction site to the *cdp-wt* gene. A gene coding CDP-CS was designed by substituting all 11 cysteine residues with serine residues, based on CDP from *A. thermocellus* strain YM4 (GenBank: AB061316.1). None of the 11 cysteine residues in CDP-WT is considered to form a disulfide bond, based on the published molecular structure [23]. In this study, we designed *cdp-cs* gene with 11 mutations that substitute free cysteines for serines and synthesized the gene to reduce the burden of mutagenesis works. The designed gene also encodes a 6×His-tag at the C-terminal of the product, and was codon-optimized for *E. coli.* Codon optimization was for effective protein expression. It was synthesized and inserted into the *Nde*I and *Xho*I restriction sites of the pET-29b vector by GenScript (Piscataway NJ, USA). Then, N terminal residues of CDP-CS were MIT (atgatcacc).

The vectors were transformed into *E. coli* BL21 (DE3), and CDP-WT and CDP-CS were expressed in Erlenmeyer flasks filled with 1 L Overnight Express auto-induction medium at 30 ˚C. After 18 h of cultivation, the cells were collected by centrifugation, and the crude enzymes were obtained after cell disruption with BugBuster reagents. The crude enzymes were purified on a TALON his-tag cobalt affinity column (Clontech Takara Bio USA Inc., San Jose CA, USA). The His-tagged target proteins were eluted with a linear gradient elution of 20 mM Tris-HCl buffer pH 7.5 with 100 mM NaCl and 500 mM imidazole. The purified His-tagged proteins were then dialyzed against 20 mM Tris-HCl buffer pH 7.5 using Amicon® Stirred Cells with 10,000 MWCO Biomax® membrane filters (Merck). Anion exchange chromatography with TOYOPEARL DEAE-650S (Tosoh Corporation, Tokyo, Japan) was employed for further purification. Highly purified CDP-WT and CDP-CS were eluted with a linear gradient of 0−250 mM NaCl in 20 mM Tris-HCl buffer pH 7.5 and were used for the following experiments. Both CDP-WT and CDP-CS were eluted when conductivity reached 7.5−9.0. Cell lysis and purification were performed within two days. The purified enzymes were stored in the elution buffer from anion exchange chromatography at 4 ˚C in 15 mL CELLSTAR® test tubes with screw caps (Greiner Bio-One International GmbH, Kremsmünster, Austria).

Activity and stability assays. For activity assay and determining kinetic parameters, the purification fractions of CDP-WT and CDP-CS with a protein concentration of 0.4 mg/mL and conductivity of 7.5−9.0 mS/cm were used and stored at 4 ˚C. This storage conditions were maintained during whole experiments. Reverse phosphorolysis activities of CDP-WT and CDP-CS were measured using 10 mM cellobiose and 10 mM αG1P in 20 mM Tris-HCl buffer pH 7.5 at 37 ˚C. The concentration of liberated phosphate ion was measured using the Malachite Green Phosphate Assay Kit (BioAssay Systems, Hayward CA, USA), and the initial rate of cellulose synthesis was determined. Enzymes were stored at 4 ˚C and activity assay was repeated every two weeks. Kinetic parameters of CDP-WT and CDP-CS were determined using 10 mM αG1P in 20 mM Tris-HCl buffer pH 7.5 at 37 ˚C by varying cellobiose concentration.

The particle sizes of CDP-WT and CDP-CS were measured at 0.6 mg/mL concentration using a Delsa Nano (Beckman Coulter, Brea, CA, USA). The storage condition was the same as the sample used in activity assay. No filtration and concentration were performed prior to the particle size measurement. Thermal stability was measured by means of thermal shift assay in 50 mM HEPES-NaOH buffer pH 7.5 using SYPRO Orange (Thermo Fisher Scientific Inc., Waltham, MA, USA) and a MX3000P real-time PCR apparatus (Stratagene Corporation, La Jolla, CA, USA), according to the protocols provided by the manufacturers. Melting temperature (T_m) was determined as the temperature at

which 50 % of the sample protein was denatured.

Crystallization and structural analysis of CDP-CS. CDP-CS was concentrated to 10 mg/ml using a Vivaspin 500 equipped with a 10 kDa cutoff membrane (Sartorius Stedim Biotech GmbH, Göttingen, Germany). To identify suitable crystallization conditions, the JCSG + suite crystallization screen (NeXtal Biotechnologies, Holland, OH, USA) was set up in 96-well plates. Aliquots of 1.0 µL protein solution were mixed with 1.0 µL well solution and the mixtures were stored at 20 ˚C for one week. Radial protein crystals were found in 100 mM sodium acetate buffer pH 4.6 with 8 % (w/v) PEG 4000. Ball-shaped crystals were obtained when 5.0 µL of 10 mg/mL protein solution was mixed with 5 µL of an optimized solution consisting of 100 mM pH 5.0 sodium acetate buffer, 12 % PEG 4000, and 5.0 % glycerol with streak seeding. Crystallization was performed according to the sitting-drop vapor diffusion method, and crystals appeared within three days. Crystals were cryocooled in liquid nitrogen with 30 % glycerol as a cryoprotectant. To generate the CDP-CS complex with αG1P, we added 1.0 µL of 10 mM αG1P adjusted to pH 5.3 with HCl to the crystal-containing droplet and incubated for 30 minutes prior to cryocooling. To generate the CDP-CS complex with cellotriose and sulfate ion, we added 1.0 µL of 10 mM cellotriose and 10 mM $(NH_4)_2SO_4$ to the crystal-containing droplet and incubated for 30 min prior to cryocooling.

X-Ray diffraction images were collected on BL-5A at the Photon Factory of the High Energy Accelerator Research Organization (KEK), using a Pilatus3 S6M detector. The diffraction images were processed with XDS [25] and Aimless in CCP4i2 [26]. A molecular model of CDP-CS was predicted with AlphaFold2 [27], and the structure was determined by molecular replacement with Phaser-MR in the Phenix suite [28]. Refinement was conducted by Phenix. refine in the Phenix suite and Coot [29]. Ligands were introduced with Coot. Ensemble refinement [30] in the Phenix suite was performed on the crystal structures of CDP-WT (PDB: 5NZ7 [23]) and CDP-CS to identify changes in structural fluctuations of mutated residues. The R package Bio3d [31] was used for trajectory analysis. Pyranose conformation was determined by Privateer in CCP4i2 [32].

RESULTS AND DISCUSSION

Activity and stability assays.

Free cysteine residues of CDP-WT are easily oxidized in aerobic industrial applications, resulting in loss of activity within a few weeks after expression and purification. Thus, improving the stability of CDP-WT is of great importance to increase catalytic efficiency and reduce costs. In this study, therefore, we designed CDP-CS, in which all 11 free cysteine residues of CDP-WT are replaced with serine (Fig. 1A), and compared the activity and stability of CDP-WT and CDP-CS over a period of six weeks after expression and purification. CDP-CS retained activity comparable to the initial activity of CDP-WT for at least six weeks, while CDP-WT had lost 80 % of its activity at six weeks (Fig. 1B). The apparent reaction parameters for the reverse reactions of CDP-WT and CDP-CS towards cellobiose are shown in Table 1. The values of CDP-CS were at the same scale as those of

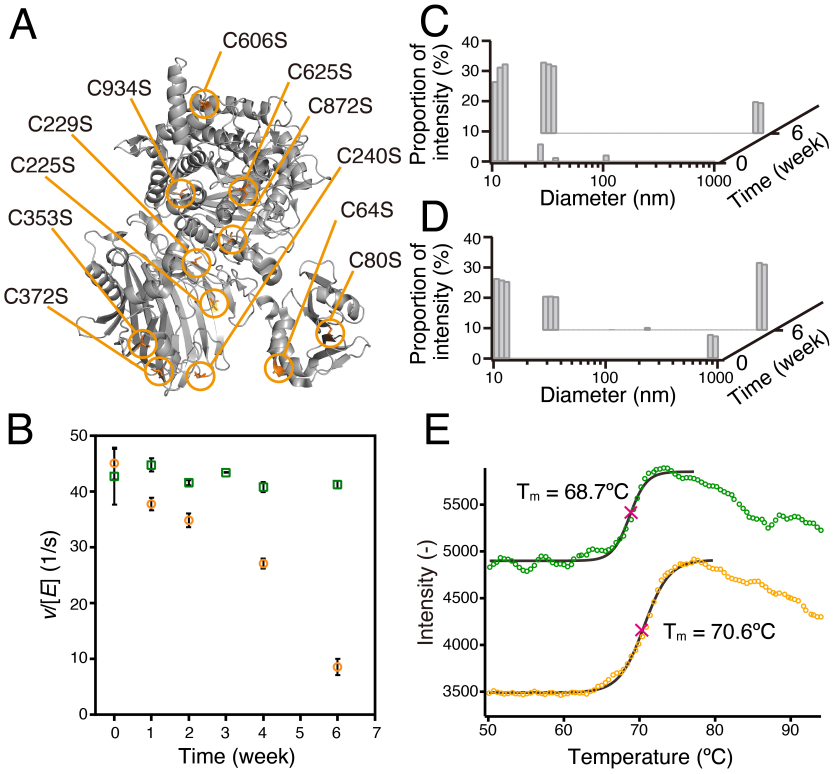

Fig. 1. The eleven cysteine-to-serine mutations in CDP-CS.

(A) The 11 mutation sites of CDP-CS. (B) Stability of CDP-CS (green squares) and CDP-WT (orange circles) during storage for six weeks. (C) DLS showed that aggregation of CDP-CS remained at a low level six weeks after expression and purification, (D) DLS showed increased aggregation of CDP-WT at six weeks after expression and purification. (E) Thermal shift assay of CDP-CS (green circles) and CDP-WT (orange circles). Data were fitted to sigmoidal curves (black lines).

Table 1. Apparent kinetic parameters and fitting errors of CDP-WT and CDP-CS for cellobiose in the reverse phosphorolysis reaction (synthetic reaction).

	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1} \cdot mM^{-1}$)
CDP-WT	2.52 ± 0.41	57.4 ± 3.4	22.8 ± 3.9
CDP-CS	4.84 ± 1.06	63.7 ± 5.9	13.2 ± 3.1

CDP-WT, indicating that the 11 cysteine-to-serine mutations had a negligible effect on the functional activity. K_m increased 2-fold by the mutation, showing CDP-CS had a lower affinity to substrates. Because CDP-WT had free cysteine (C625) next to catalytic acid D624, this cysteine is responsible for subsite recognition [23], although serine residue can be the weak alternative.

We found that CDP-WT lost its activity concomitantly with aggregation. To examine if the 11 cysteine-to-serine mutations affected aggregation, we measured the particle sizes of CDP-WT and CDP-CS by means of a dynamic light scattering (DLS) technique. Although aggregates of about 1,000 nm in size accounted for 65.2 % of CDP-WT at six weeks after expression and purification, the corresponding figure for CDP-CS was only 30.7 % (Figs.1C and 1D). These results suggest that intermolecular and/or non-specific disulfide bond formation involving the cysteine residues makes a substantial contribution to the aggregation of CDP-WT. Furthermore, these results showed that aggregation of CDP-WT is not the main reason of loss of activity. Aggregation of CDP-WT may occur after loss of the activity, because population of aggregation showed 65.2 % after six weeks while CDP-WT lost 80 % of the activity. Activity loss of CDP-WT can cause the change in protein fold by oxidation, which in turn cause the aggregation.

To examine how the cysteine-to-serine mutations influenced the conformational stability of CDP-WT and CDP-CS, thermal shift assay was performed in HEPES-NaOH buffer pH 7.5. The T_m of CDP-CS was 68.7 ˚C, while that of CDP-WT was 70.6 ˚C (Fig. 1E). Thus, in contrast to the improved stability to oxidation, the thermal stability of CDP-CS was slightly decreased. This is consistent with a previous finding that cysteine-to-serine mutation decreased the thermal stability of α-galactosidase [33]. However, in other cases, the elimination of free cysteines improved the thermal stability [18−20]. It has been reported that mutation of surface-accessible cysteine residues prevents aggregation and improves thermal stability, whereas mutation of hydrophobically interacting cysteine residues negatively affects the thermal stability of a mitogen-activated protein kinase, c-Jun N-terminal kinase 1 [34]. Serine-to-cysteine mutation improved the thermal stability and van der Waals interactions of β-lactamase [35], suggesting that some cysteines contribute to thermal stability by strengthening residue-residue interactions. In our study, however, mutation of some of the 11 free cysteines negatively affected the thermal stability, presumably by weakening van der Waals interactions. Predicting how specific cysteine residues contribute to oxidative and thermal stability remains a challenging goal.

Crystal structure of CDP-CS.

We were able to obtain high-quality crystals of CDP-CS for an X-ray diffraction study, due to its improved stability. [36] The molecular structure of CDP-CS was determined at 1.37 Å resolution (Table S1), while the previously determined crystal structure of CDP-WT had a resolution of 2.30 Å [23]. As is the case for CDP-WT [23], CDP-CS exists as a dimer with the active sites facing each other in the crystal (Fig. 2A). Each CDP-CS monomer contains 4 domains plus a linker: an N-terminal domain (residue 1−119), a β-sandwich domain (residue 120−419), an L-shaped linker (420−457), an $(\alpha/\alpha)_6$-barrel catalytic domain (458−903), and a C-terminal jelly roll domain (904−957). The β-sandwich domain and the $(\alpha/\alpha)_6$-barrel catalytic domain with the C-terminal jelly roll domain are well conserved in GH94 family members active towards disaccharides [37−40]. The dimeric interface of CDP-WT and CDP-CS is formed mainly by the N-terminal domains and β-sandwich domains, although the dimeric interfaces of most GH94 enzymes consist of β-sandwich domains and $(\alpha/\alpha)_6$-barrel catalytic domains [37−41]. Each CDP-CS monomer has two Cl^- ions placed near H140 and W622 in the crystal structure, as previously reported in the crystal structure of CDP-WT [23]. The 11 cysteine-to-serine mutations did not markedly alter the overall structure; the molecular structures of CDP-CS and CDP-WT showed a Cα root-mean-square deviation (RMSD) of only 0.286 Å (Fig. 2B). Among the 11 mutated residues, the greatest difference was found in C625S, adjacent to the catalytic acid D624, involving a movement of 1.63 Å away from the catalytic center (Fig. 3A).

To examine substrate recognition of cellotriose, we soaked a CDP-CS crystal in cellotriose and ammonium sulfate solution, and the molecular structure of the resulting CDP-CS complex with cellotriose was determined at 1.21 Å resolution (Table S1). The Cα RMSD between the crystal structure of CDP-CS and that of CDP-CS with cellotriose and sulfate ion was 0.16 Å; thus, the overall structures were almost identical. Two cellotriose molecules were found in the groove-like catalytic sites of CDP-CS. The two cellotriose chains had an anti-parallel orientation at a distance of 2.20 nm (Fig. 2C). These cellotrioses were accommodated in subsites +1 to +3. The occupancy was approximately 50 %, possibly because the CDP-CS crystal was soaked in cellotriose solution for only a short time.

The glucose moiety at subsite +2 was more distinct than other glucose moieties (Fig. 2D), suggesting that substrate binding at this subsite is stronger. This is consistent with the kinetic parameters for several cellooligosaccharides; the K_m for phosphorolysis was smaller when the DP was larger than 2 [8], and the K_m for synthesis was detectable when the DP of glucosyl acceptor was larger than 1 [24]. In addition, the coordination of the glucose moiety at subsite +3 was flipped compared to that in the previously solved structure of CDP-WT with cellotetraose (5NZ8), suggesting that O6 of the glucose residue at subsite +3 is not recognized or that substrate recognition at this subsite is weak, permitting rotation of the glucose residue (Fig. 3B).

Examination of the crystal structure of the CDP-CS complex with cellotriose revealed several residues involved in the recognition of cellotriose (Fig. 2D). At subsite +1, E810 interacted with O2 and O3 of the glucose moiety, and Y804 interacted with O2. At subsite +2, residues of the opposing monomer of CDP-CS interacted with the glucose moiety; E296 and D297 formed hydrogen bonds with O2 and O3 of the glucose moiety, respectively, and Y300 showed a

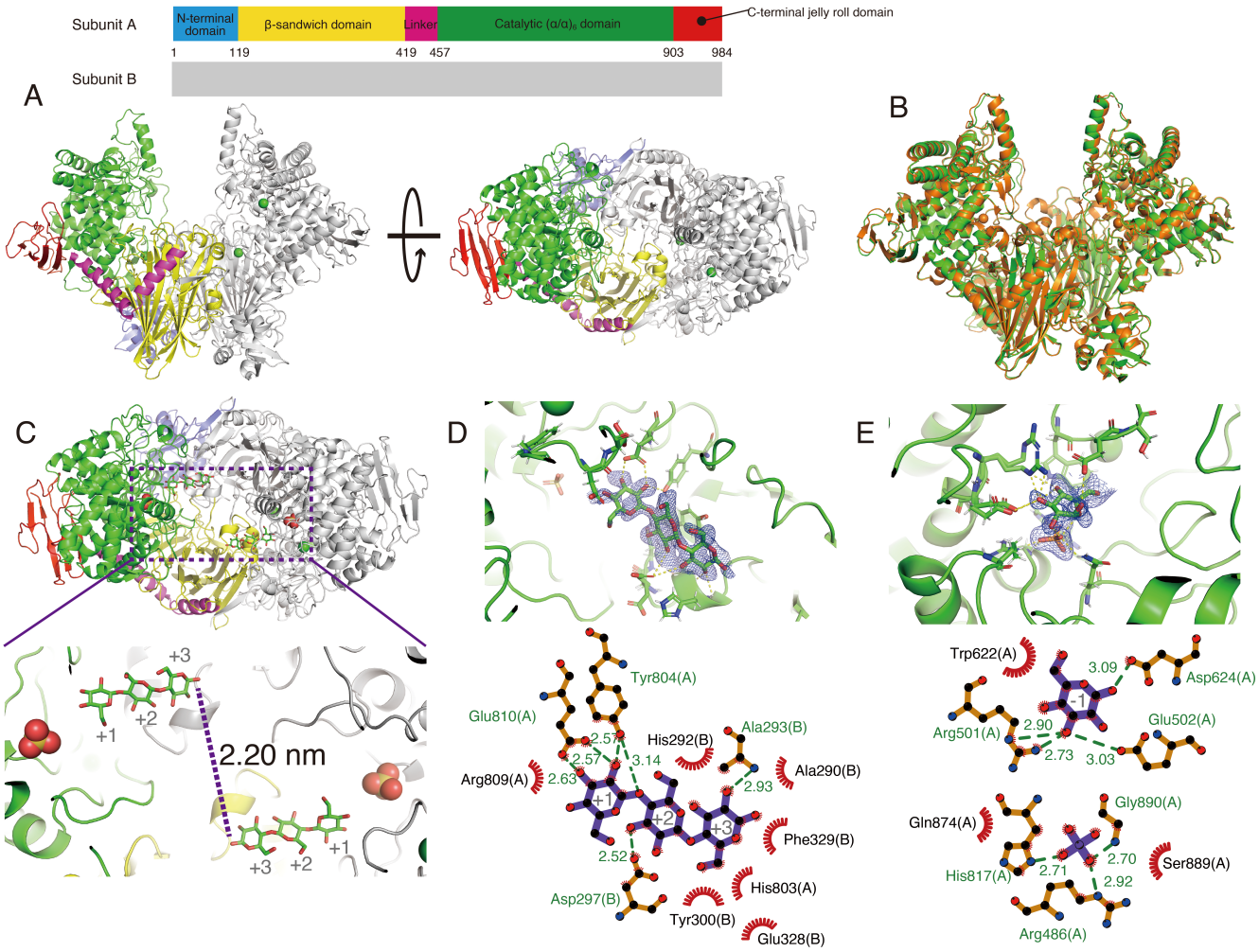

Fig. 2. X-ray crystal structure of CDP-CS.

(A) Dimeric structure of CDP-CS. The N-terminal domain, the β-sandwich domain, the linker, the catalytic $(\alpha/\alpha)_6$ domain, and C-terminal jelly roll domain are colored blue, yellow, purple, green, and red, respectively. The second monomer of the dimer is colored gray. Chloride ions are shown as green spheres. (B) The CDP-CS structure is superimposed on the previously solved CDP-WT structure (PDB ID: 5NZ7). (C) CDP-CS structure liganded with cellotriose. Cellotriose was aligned in an anti-parallel manner in the CDP-CS dimeric structure. Numbers near the ligands were the subsite numbers. (D) A polder map for cellotriose in subsites +1 to +3 of CDP-CS and a schematic illustration of cellotriose recognition. (E) A polder map for αG1P in subsite −1 of CDP-CS and a schematic illustration of αG1P recognition.

hydrophobic stacking interaction with the glucose moiety. Similarly, E296 and H292 of the opposing monomer of CDP-CS formed hydrogen bonds with O3 and O2 of the glucose moiety, respectively. Thus, substrates were recognized by hydrogen bonding at subsites +1 to +3 and hydrophobic stacking interaction at subsite +2, so that O2 and O3 of the substrates play key roles in the interaction with the enzyme.

In addition, we soaked a CDP-CS crystal in αG1P solution and solved the molecular structure of the resulting complex at 1.68 Å resolution (Table S1). The RMSD for Cα of the complex compared with CDP-CS was very small, 0.33 Å. In the catalytic site, glucose and phosphate ion of αG1P was found separately, possibly showing a transition state of the protein upon αG1P binding. The glucose moiety of αG1P was found at subsite −1, with approximately 50 % occupancy (Fig. 2E), and it was in $^{1}S_{3}$ conformation, in contrast to the $^{4}C_{1}$ conformation of the glucose moiety at subsite −1 in the presence of phosphate ion in a previously solved crystal structure of CDP-WT (5NZ8) [23] (Fig. 3C). The glucose moiety of αG1P was separate from the phosphate ion, whose occupancy was approximately 30 %. We found that E502 of CDP-CS formed a hydrogen bond with O3 of the glucose moiety of αG1P and that in one conformation, R501 was located near O3 and O4 of the glucose moiety (Fig. 2E). Note that R501 did not have multiple conformations in the absence of the ligand. We found that several residues of CDP-CS are involved in recognizing the phosphate ion at the catalytic site: for example, R486, H817, Q874, and G890 formed hydrogen bonds with the phosphate ion. These residues are highly conserved among GH94 phosphorylases, as previously discussed [42]. The locations of these residues are consistent with those in the X-ray crystal structure of CDP-WT [23], except for C625 (S625 for CDP-CS), whose Cα is displaced by 1.63 Å away from the catalytic center (Fig. 3A). When αG1P or cellotriose was bound to CDP-CS, S625 of CDP-CS showed 1.90 Å or 2.31 Å deviation, respectively (Table 2, Figs. 3B and C).

Next, we examined the deviations of the residues around the catalytic site of CDP-CS bound to different substrates (Table 2, Fig. 3D). Although the overall structures were almost identical to that of CDP-CS without any ligand (RMSD < 0.4 Å), residues around the catalytic site showed relatively large changes in coordinates. As shown in Table 2, the movements of residues around the catalytic site depended upon the ligand: the coordinates of D624 and S625 showed no change when cellotriose was accommodated in

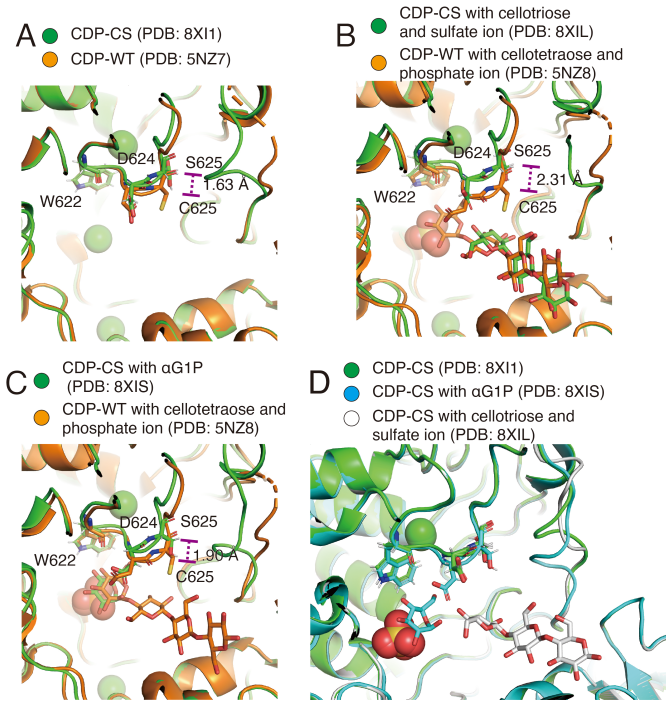

Fig. 3. Coordinates of residues around the catalytic site of CDP-CS.
(A) Structural deviations in residues between CDP-CS and CDP-WT [23]. Chloride ions are shown as green spheres. (B) Structural deviations in residues between CDP-CS and CDP-WT with cellodextrin as liganded. (C) Structural deviations in residues of CDP-CS with G1P as liganded. (D) The structure of CDP-CS is superimposed on the structure of CDP-CS with cellotriose and sulfate ion and that of CDP-CS with αG1P.

Table 2. Deviations of Cα of the residues around the catalytic site upon substrate docking.

	CDP-WT			CDP-CS		
	W622	D624	C625	W622	D624	S625
CDP-CS	0.212	0.870	1.63	–	–	–
CDP-CS with αG1P	0.530	1.09	1.90	0.220	1.02	0.743
CDP-CS with cellotriose and SO_4^{2-}	0.751	2.01	2.31	0.0661	0.134	0.135

Two CDP-WT structures were used; 5NZ7 for comparison with CDP-CS, and 5NZ8 for comparison with CDP-CS in the presence of αG1P, and CDP-CS in the presence of cellotriose and SO_4^{2-}. The units are Å.

subsites +1 to +3. However, D624 and S625 moved 1.02 Å and 0.743 Å towards the catalytic site, respectively, when αG1P was accommodated in subsite −1. In both cases, phosphate ion or sulfate ion was bound to the phosphate site; thus, binding of a glucose moiety at subsite −1 might induce the movement of D624 and S625. Similarly, a previous comparison of the structures of CDP-WT and CDP-WT with phosphate ion and cellotetraose at subsites −1 to +3 found that D624 and C625 moved 1.43 Å and 0.708 Å upon substrate binding [23].

Ensemble refinement of the crystal structures of CDP-WT and CDP-CS.

Next, we performed ensemble refinement of the crystal structures of CDP-WT (Fig. 4A) and CDP-CS (Fig. 4B) to examine the structural fluctuations of the mutated residues. Root-mean-square fluctuation (RMSF) of Cα of each residue was quantified from the structure obtained by ensemble refinement (Fig. 4C). Overall, we found that CDP-WT and CDP-CS had similar fluctuations: the fluctuations of the N-terminal domain and the loop of residues 789−802 were larger than those of the other domains. Mutation of the 11 free cysteine residues to serine decreased the average RMSF from 0.344 (CDP-WT) to 0.215 (CDP-CS), suggesting that the elimination of free cysteine decreased the Cα vibrational amplitude. Among the mutated residues (Fig. 4D and Table S2), C225S and C934S were fixed and showed small fluctuations. The mutations C229S and C872S also caused no significant residue fluctuations. The C64S, C80S, C240S, C353S, and C606S mutations resulted in increased side chain rotation, suggesting that residue-residue interactions were impaired. The effect of the C80S mutation was especially large (Fig. 4D and Table S3). These impaired residue-residue interactions could be the reason for the decreased thermal stability of CDP-CS, as found in the thermal shift assay (Fig. 1E). On the other hand, the C372S and C625S mutations decreased the RMSF values from 1.21 to 0.441 and from 1.31 to 0.520, respectively. These cysteine residues are exposed to the solution, and thus could be vulnerable to oxidation. Indeed, C625 is located next to the catalytic acid D624, and thus could contribute to the instability of CDP-WT to oxidation. Previous studies have found that some free cysteine residues help to maintain protein folding by strengthening residue-residue hydrophobic

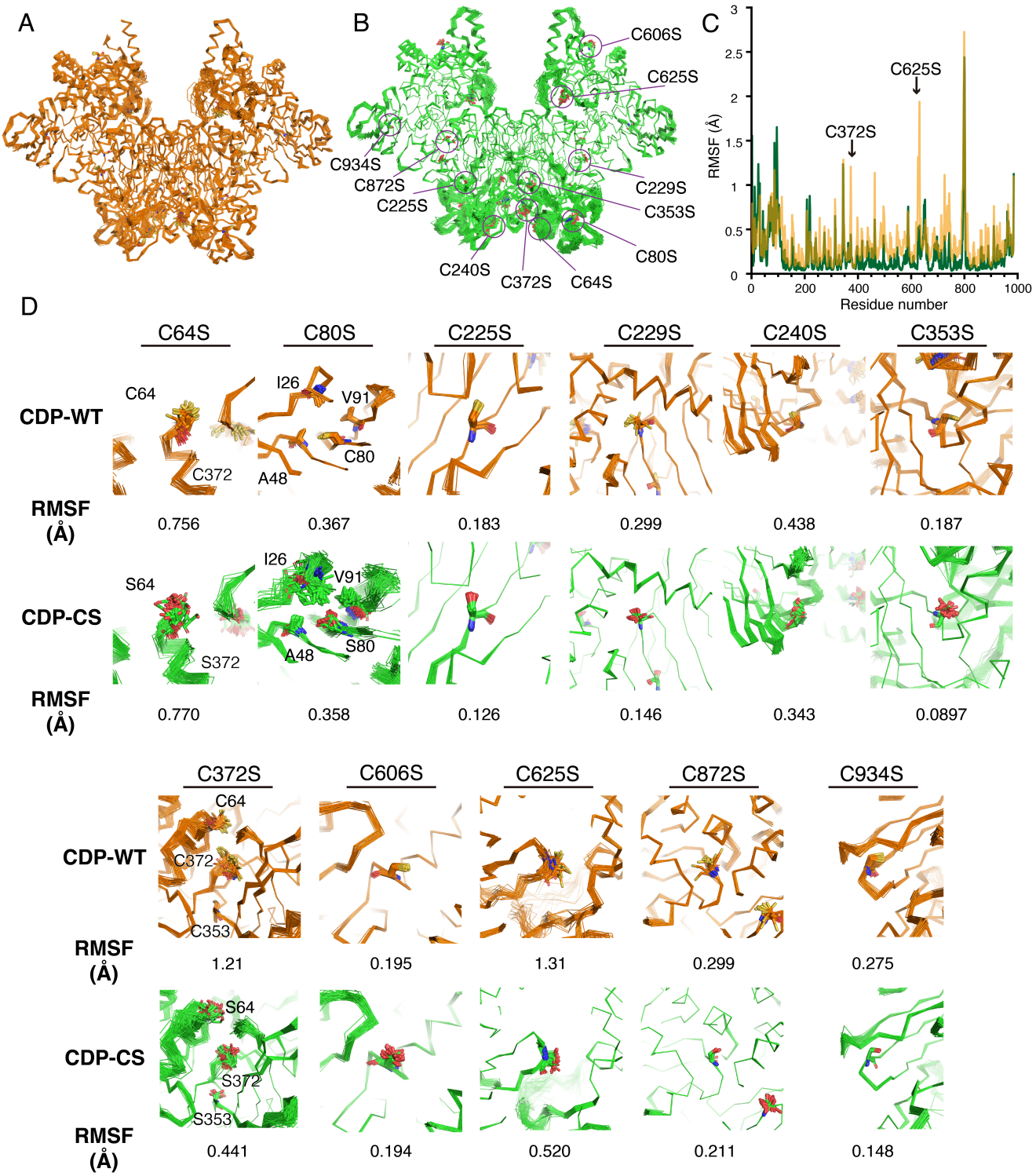

Fig. 4. Ensemble refinement of the crystal structures of CDP-WT and CDP-CS.
(A) Ensemble refinement models of CDP-WT. (B) Ensemble refinement models of CDP-CS. (C) Change in main chain fluctuation. Averaged RMSF values of the two monomers in the dimer are shown. CDP-WT and CDP-CS are shown in orange and green, respectively. (D) The fluctuation of the main chain around mutated residues in CDP-WT and CDP-CS.

interaction [43], whereas others are easily damaged by oxidation [34, 44, 45]. The results of our ensemble refinement highlight the differences among cysteine residues, and suggest that ensemble refinement may be an effective method for evaluating the effects of specific mutations, as a previous study suggested [46].

Free cysteine residues in GH94 enzymes.

A. thermocellus YM4 is an anaerobic bacterium originally found in volcanic soil at Izu peninsula, Japan [47]. The sulfur-rich growth environment may create a reductive environment inside the cell [48], and this may account for stability of the free-cysteine-rich enzyme *in vivo*, in contrast to the *in vitro* situation. Many other GH94 enzymes are intracellular phosphorylases isolated from anaerobic bacteria, and have potential industrial applications based on the reversibility of the phosphorolysis reaction [49, 50]. Interestingly, they also contain many free cysteine residues, as predicted by Alphafold2 [27] (Table 3). Thus, the approach we adopted in this study to stabilize CDP-WT could be

Table 3. Numbers of cysteines in enzymes belonging to the GH94 family.

Protein	Organism	Anaerobic/aerobic	Uniprot	Number of free cysteines	Total sequence length
Cellobionic acid phosphorylase	*Saccharophagus degradans* (strain 2-40 / ATCC 43961 / DSM 17024)	Aerobic	Q21MB1	12	788
Cellobiose phosphorylase	*Cellulomonas gilvus*	Anaerobic	O66264	5	822
Cellobiose phosphorylase	*Cellulomonas uda*	Anaerobic	Q7WTR65	5	822
Cellobiose phosphorylase	*Ruminococcus albus* 7 = DSM 20455	Anaerobic	E0LCW6	10	827
Cellobiose phosphorylase	*Thermoclostridium stercorarium* (*Clostridium stercorarium*)	Anaerobic	P77846	20	780
Cellobiose phosphorylase	*Acetivibrio thermocellus* (*Clostridium thermocellum*) YM4	Anaerobic	Q8VP44	11	811
β-1,2-oligoglucan phosphorylase	*Lachnoclostridium phytofermentans* ISDg	Anaerobic	A9KJS6	14	1113
Laminaribiose phosphorylase	*Paenibacillus* sp. YM1	Anaerobic	D7UT17	4	911
Chitobiose phosphorylase	*Vibrio proteolyticus*	Anaerobic	Q76IQ9	13	801

applicable to other GH94 enzymes containing free cysteine residues.

CONCLUSION

In this study, we demonstrated that the substitution of the 11 free cysteine residues of CDP-WT by mutation to serine improved the stability of CDP-WT to oxidation. Ensemble refinement of the crystal structures of CDP-WT and CDP-CS identified two cysteine residues, namely C372 and C625, associated with susceptibility to oxidation and one cysteine residue (C80) important for maintaining the protein folding. Since other GH94 enzymes also have many free cysteine residues, our results may be helpful for improving the stability of other enzymes, as well as CDP-WT, thereby reducing the cost of using such enzymes in commercial catalytic reactions.

CONFLICTS OF INTERESTS

The authors declare that they have no competing interests.

ACKNOWLEDGMENTS

We thank Prof. Motomitsu Kitaoka of Niigata University for providing the pET-28a vector harboring the *cdp-wt* gene. We thank Prof. Tsuguyuki Saito and Assist. Prof. Shuji Fujisawa of the University of Tokyo for DLS measurement. We also thank Prof. Shinya Fushinobu for helpful discussions. This work was supported by the "Development of recycling system for bio-organic materials in space" program, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Coordination Funds for Promoting AeroSpace Utilization, Japan; Grant Number JPJ000959. This work was also supported by JSPS KAKENHI (Grant no. 22J12566 to TK Grant no. 19K15884 to NS) and by a Grant-in-Aid for Innovative Areas from MEXT (Grant no. 18H05494 to KI), the Academy of Finland through research grant SA-FOSSOK [Decision No. 309384]. KI thanks the Finnish Funding Agency for Innovation for the support of the Finland Distinguished Professor Program "Advanced approaches for enzymatic biomass utilization and modification (BioAD)"

REFERENCES

[1] Drula E, Garron M-L, Dogan S, Lombard V, Henrissat B, Terrapon N. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res. 2022; 50: D571−7.

[2] Sheth K, Alexander JK. Cellodextrin phosphorylase from *Clostridium thermocellum.* Biochim Biophys Acta - Gen Subj. 1967; 148: 808−10.

[3] Raman B, McKeown CK, Rodriguez M, Brown SD, Mielenz JR. Transcriptomic analysis of *Clostridium thermocellum* ATCC 27405 cellulose fermentation. BMC Microbiol. 2011; 11: 1−15.

[4] Zhang YHP, Lynd LR. Cellulose utilization by *Clostridium thermocellum*: Bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci USA. 2005; 102: 7321−7325.

[5] Lou J, Dawson KA, Strobel HJ. Cellobiose and cellodextrin metabolism by the ruminal bacterium *Ruminococcus albus*. Curr Microbiol. 1997; 35: 221−7.

[6] Alexander JK. Cellodextrin phosphorylase from *Clostridium thermocellum.* in *Methods in Enzymology*, New York: Academic Press; 1972. p. 948−53.

[7] Samain E, Lancelon-Pin C, Férigo F, Moreau V, Chanzy H, Heyraud A, et al. Phosphorolytic synthesis of cellodextrins. Carbohydr Res. 1995; 271: 217−26.

[8] Hiraishi M, Igarashi K, Kimura S, Wada M, Kitaoka M, Samejima M. Synthesis of highly ordered cellulose II *in vitro* using cellodextrin phosphorylase. Carbohydr Res. 2009; 344: 2468−73.

[9] Pylkkänen R, Mohammadi P, Arola S, De Ruijter JC, Sunagawa N, Igarashi K, et al. *In vitro* synthesis and self-assembly of cellulose II nanofibrils catalyzed by the reverse reaction of *Clostridium thermocellum* cellodextrin phosphorylase. Biomacromolecules. 2020; 21: 4355−64.

[10] Kuga T, Sunagawa N, Igarashi K. Enzymatic synthesis of cellulose in space: gravity is a crucial factor for building cellulose II gel structure. Cellulose. 2022; 29: 2999−3015.

[11] Serizawa T, Fukaya Y, Sawada T. Self-assembly of cellulose oligomers into nanoribbon network structures based on kinetic control of enzymatic oligomerization. Langmuir. 2017; 33: 13415−22.

[12] Shintate K, Kitaoka M, Kim Y-K, Hayashi K. Enzymatic synthesis of a library of β-(1→4) hetero- D-glucose and D-xylose-based oligosaccharides employing cellodextrin phosphorylase. Carbohydr Res. 2003; 338: 1981−90.

[13] Yataka Y, Sawada T, Serizawa T. Multidimensional self-assembled structures of alkylated cellulose oligomers synthesized via *in vitro* enzymatic reactions. Langmuir, 2016; 32: 10120−5.

[14] Yataka Y, Sawada T, Serizawa T. Enzymatic synthesis and post-functionalization of two-dimensional crystalline cellulose oligomers with surface-reactive groups. Chem Commun. 2015; 51: 12525−8.

[15] Sheth K, Alexander JK. Purification and properties of β-1,4-oligoglucan:orthophosphate glucosyltransferase from *Clostridium thermocellum*. J Biol Chem. 1969; 244: 457−64.

[16] Lacy ER, Baker M, Brigham-Burke M. Free sulfhydryl measurement as an indicator of antibody stability. Anal Biochem. 2008; 382: 66−68.

[17] Buchanan A, Clementel V, Woods R, Harn N, Bowen MA, Mo W, et al. Engineering a therapeutic IgG molecule to address cysteinylation, aggregation and enhance thermal stability and expression. mAbs. 2013; 5: 255−62.

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

[19] 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.

[20] Li J, Zhang S, Yi Z, Pei X, Wu Z. Removal of the free cysteine residue reduces irreversible thermal inactivation of feruloyl esterase: evidence from circular dichroism and fluorescence spectra. Acta Biochim Biophys Sin. 2015; 47: 612−19.

[21] Amaki Y, Nakano H, Yamane T. Role of cysteine residues in esterase from *Bacillus stearothermophilus* and increasing its thermostability by the replacement of cysteines. Appl Microbiol Biotechnol. 1994; 40: 664−8.

[22] Zheng J, Yang T, Zhou J, Xu M, Zhang X, Rao Z. Elimination of a free cysteine by creation of a disulfide bond increases the activity and stability of *Candida boidinii* formate dehydrogenase. Appl Environ Microbiol. 2017; 83: e02624−16.

[23] O'Neill EC, Pergolizzi G, Stevenson CEM, Lawson DM, Nepogodiev SA, Field RA. Cellodextrin phosphorylase from *Ruminiclostridium thermocellum:* X-ray crystal structure and substrate specificity analysis. Carbohydr Res. 2017; 451: 1−15.

[24] Krishnareddy M, Kim YK, Kitaoka M, Mori Y, Hayashi K. Cellodextrin phosphorylase from *Clostridium thermocellum* YM4 strain expressed in *Escherichia coli.* J Appl Glycosci. 2002; 49: 1−8.

[25] Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010; 66: 125−32.

[26] Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011; 67: 235−42.

[27] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al.: Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 596: 583−9.

[28] Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, et al.: Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Biol Crystallogr. 2019; 75: 861−77.

[29] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60: 2126−32.

[30] Burnley BT, Afonine PV, Adams PD, Gros P. Modelling dynamics in protein crystal structures by ensemble refinement. eLife. 2012; 1: e00311.

[31] Grant BJ, Skjærven L, Yao XQ. The Bio3D packages for structural bioinformatics. Protein Sci. 2021; 30: 20−30.

[32] Agirre J, Iglesias-Fernández J, Rovira C, Davies GJ, Wilson KS, Cowtan KD. Privateer: software for the conformational validation of carbohydrate structures. Nat Struct Mol Biol. 2015; 22: 833−34.

[33] Qiu H, Honey DM, Kingsbury JS, Park A, Boudanova E, Wei RR, et al. Impact of cysteine variants on the structure, activity, and stability of recombinant human α-galactosidase A. Protein Sci. 2015; 24: 1401−11.

[34] Nakaniwa T, Fukada H, Inoue T, Gouda M, Nakai R, Kirii Y, et al. Seven cysteine-deficient mutants depict the interplay between thermal and chemical stabilities of individual cysteine residues in mitogen-activated protein kinase c-Jun N-terminal kinase 1. Biochemistry. 2012; 51: 8410−21.

[35] Santos J, Risso VA, Sica MP, Ermácora MR. Effects of serine-to-cysteine mutations on β-lactamase folding. Biophys J. 2007; 93: 1707−18.

[36] Deller MC, Kong L, Rupp B. Protein stability: a crystallographer's perspective. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2016; 72: 72−95.

[37] Hidaka M, Honda Y, Kitaoka M, Nirasawa S, Hayashi K, Wakagi T, et al. Chitobiose Phosphorylase from *Vibrio proteolyticus,* a member of glycosyl transferase family 36, has a clan GH-L-like $(\alpha/\alpha)_6$ barrel fold. Structure. 2004; 12: 937−47.

[38] Hidaka M, Kitaoka M, Hayashi K, Wakagi T, Shoun H, Fushinobu S. Structural dissection of the reaction mechanism of cellobiose phosphorylase. Biochem J. 2006; 398: 37−43.

[39] Kuhaudomlarp S, Walpole S, Stevenson CEM, Nepogodiev SA, Lawson DM, Angulo J, et al. Unravelling the specificity of laminaribiose phosphorylase from *Paenibacillus* sp. YM-1 towards donor substrates glucose/mannose 1-phosphate by using X-ray crystallography and saturation Transfer Difference NMR Spectroscopy. ChemBioChem. 2019; 20: 181−92.

[40] Nam YW, Nihira T, Arakawa T, Saito Y, Kitaoka M, Nakai H, et al. Crystal structure and substrate recognition of cellobionic acid phosphorylase, which plays a key role in oxidative cellulose degradation by microbes. J Biol Chem. 2015; 290: 18281−92.

[41] Van Hoorebeke A, Stout J, Kyndt J, De Groeve M, Dix I, Desmet T, et al. Crystallization and X-ray diffraction studies of cellobiose phosphorylase from *Cellulomonas uda*. Acta Crystallogr F Struct Biol Commun. 2010; 66: 346−51.

[42] Saburi W, Nihira T, Nakai H, Kitaoka M, Mori H. Discovery of solabiose phosphorylase and its application for enzymatic synthesis of solabiose from sucrose and lactose.

Sci Rep. 2022; 12: 1−11.
[43] Sandgren M, Gualfetti PJ, Paech C, Paech S, Shaw A, Gross LS, et al. The *Humicola grisea* Cel12A enzyme structure at 1.2 Å resolution and the impact of its free cysteine residues on thermal stability. Protein Sci. 2003; 12: 2782−93.
[44] Hamza MA, Engel PC. Enhancing long-term thermal stability in mesophilic glutamate dehydrogenase from *Clostridium symbiosum* by eliminating cysteine residues. Enzyme Microb Technol. 2007; 41: 706−10.
[45] Jiao L, Chi H, Xia B, Lu Z, Bie X, Zhao H, et al. Thermostability improvement of L-asparaginase from *Acinetobacter soli* via consensus-designed cysteine residue substitution. Molecules. 2022; 27: 6670.
[46] Matsuyama K, Kishine N, Fujimoto Z, Sunagawa N, Kotake T, Tsumuraya Y, et al. Unique active-site and subsite features in the arabinogalactan-degrading GH43 exo-β-1,3-galactanase from *Phanerochaete chrysosporium.* J Biol Chem. 2020; 295: 18539−52.
[47] Mori Y. Characterization of a symbiotic coculture of *Clostridium thermohydrosulfuricum* YM3 and *Clostridium thermocellum* YM4. Appl Environ Microbiol. 1990; 56: 37−42.
[48] Ulrich K, Jakob U. The role of thiols in antioxidant systems. Free Radic Biol Med. 2019; 140: 14−27.
[49] Kitaoka M. Diversity of phosphorylases in glycoside hydrolase families. Appl Microbiol Biotechnol. 2015; 99: 8377−90.
[50] O'Neill EC, Field RA. Enzymatic synthesis using glycoside phosphorylases. Carbohydr Res. 2015; 403: 23−37.