Appl Microbiol Biotechnol (2012) 95:661–669
DOI 10.1007/s00253-011-3791-4

Enhanced activity of *Thermotoga maritima* cellulase 12A by mutating a unique surface loop

Ya-Shan Cheng · Tzu-Ping Ko · Jian-Wen Huang · Tzu-Hui Wu · Cheng-Yen Lin · Wenhua Luo · Qian Li · Yanhe Ma · Chun-Hsiang Huang · Andrew H.-J. Wang · Je-Ruei Liu · Rey-Ting Guo

Received: 18 October 2011 / Revised: 20 November 2011 / Accepted: 23 November 2011 / Published online: 15 December 2011

Abstract Cellulase 12A from *Thermotoga maritima* (*Tm*Cel12A) is a hyperthermostable β-1,4-endoglucanase. We recently determined the crystal structures of *Tm*Cel12A and its complexes with oligosaccharides. Here, by using site-directed mutagenesis, the role played by Arg60 and Tyr61 in a unique surface loop of *Tm*Cel12A was investigated. The results are consistent with the previously observed hydrogen bonding and stacking interactions between these two residues and the substrate. Interestingly, the mutant Y61G had the highest activity when compared with the wild-type enzyme and the other mutants. It also shows a wider range of working temperatures than does the wild type, along with retention of the hyperthermostability. The k_{cat} and K_m values of Y61G are both higher than those of the wild type. In conjunction with the crystal structure of Y61G–substrate complex, the kinetic data suggest that the higher endoglucanase activity is probably due to facile dissociation of the cleaved sugar moiety at the reducing end. Additional crystallographic analyses indicate that the insertion and deletion mutations at the Tyr61 site did not affect the overall protein structure, but local perturbations might diminish the substrate-binding strength. It is likely that the catalytic efficiency of *Tm*Cel12A is a subtle balance between substrate binding and product release. The activity enhancement by the single mutation of Y61G provides a good example of engineered enzyme for industrial application.

Keywords Thermostability · Endoglucanase · Crystal structure · Protein engineering · Enzyme kinetics

Introduction

Cellulose, an important biomass resource, is one of the major components in the plant cell wall, constituting approximately

Y.-S. Cheng and T.-P. Ko contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00253-011-3791-4) contains supplementary material, which is available to authorized users.

Y.-S. Cheng · T.-H. Wu · J.-R. Liu (✉)
Institute of Biotechnology, National Taiwan University,
Taipei 106, Taiwan
e-mail: jrliu@ntu.edu.tw

T.-P. Ko · A. H.-J. Wang
Institute of Biological Chemistry, Academia Sinica,
Taipei 115, Taiwan

J.-W. Huang · C.-Y. Lin
Genozyme Biotechnology Inc.,
Taipei 114, Taiwan

J.-W. Huang · C.-Y. Lin
AsiaPac Biotechnology Co., Ltd.,
Dongguan 523808, China

W. Luo
College of Food Science, South China Agricultural University,
Guangzhou 510642, China

Q. Li · Y. Ma · C.-H. Huang · R.-T. Guo (✉)
Industrial Enzymes National Engineering Laboratory,
Tianjin Institute of Industrial Biotechnology,
Chinese Academy of Sciences,
Tianjin 300308, China
e-mail: guo_rt@tib.cas.cn

J.-R. Liu
Department of Animal Science and Technology,
National Taiwan University,
Taipei 106, Taiwan

35% to 50% of plant dry weight (Jarvis 2003; Lynd et al. 2002). Cellulose is composed of glucose units linked by β-1,4-glycosidic bond. The resulting polysaccharide chains organize tightly together to form crystalline architectures that resist destructing forces from the elements. On the other hand, many herbivores and microbes degrade the plant matter to extract soluble sugars as an energy source by various enzymes including cellulase, xylanase, and mannanase (Selinger et al. 1996). The catalytic mechanism of cellulase involves hydrolyzing the β-1,4-glycosidic bond by nucleophilic attacks and acid–base reactions (Birsan et al. 1998; Vocadlo and Davies 2008). Cellulases can be divided into three general groups: endoglucanase, cellobiohydrolase, and β-glucosidase (Henrissat et al. 1998). Endoglucanase randomly cleaves cellulose into fragments; cellobiohydrolase processively acts on the chain termini to release cellobiose, and β-glucosidase hydrolyzes cellobiose into glucose.

Widespread applications of cellulases have been found in the food industries, the processing of feed and textile, as well as biofuel production (Bhat 2000; Ohmiya et al. 1997). For a long time, however, the use of cellulases obtained from mesophiles (e.g., *Trichoderma reesei*) suffered from low thermostability (Sandgren et al. 2003b). Using thermostable cellulases can increase the production efficiency in industries that prefer high-temperature reaction conditions, such as brewing and bioethanol production (Blumer-Schuette et al. 2008; Szijártó et al. 2011; Vieille and Zeikus 2001; Viikari et al. 2007). To exploit the high catalytic efficiency at elevated temperatures, scientists in both academic and industrial organizations have searched for suitable enzymes by screening the unknown in nature or modifying the known in the laboratory (Dumon et al. 2008; Dutta et al. 2008; Hua et al. 2010; Kapoor et al. 2008; Sandgren et al. 2003b; Yang et al. 2010). In general, there are two strategies of enzyme modification: (1) directed evolution by randomly mutating the enzyme-encoding gene and subsequent selection for desirable properties and (2) rational engineering by specifically mutating the gene, based on the structural information of the enzyme (Bornscheuer and Pohl 2001; Böttcher and Bornscheuer 2010; Percival-Zhang et al. 2006; Schülein 2000). Furthermore, cellulases isolated from hyperthermophiles such as *Pyrococcus furiosus* and *Thermotoga maritima* often lack sufficient production level and specific activity (Blumer-Schuette et al. 2008; Viikari et al. 2007). Therefore, conservation of activity is also a key point when modifying the enzymes for industrial use.

Some recently determined crystal structures of the above-mentioned thermostable cellulases allowed further understanding of the catalytic mechanism and specific modifications of the proteins (Crennell et al. 2006; Ilari et al. 2009; Yang et al. 2008). Even so, few industrial cellulases with both high thermostability and high enzyme activity have been obtained. In our previous study of the crystal structures of *T. maritima* cellulase 12A (*Tm*Cel12A) and its complex with oligosaccharide, a unique surface loop between the strands A3 and B3 was found to protrude over the active-site cleft, forming an enclosure around the −1 sugar of the substrate (Cheng et al. 2011). This loop, containing Arg60 and Tyr61, is longer than its equivalents in the other GH12-family enzymes and deviates significantly from them in sequence and structure. Arg60 interacts with three sugar units in the −2, −1, and +1 subsites by hydrogen bonding, while Tyr61 stacks with the +2 sugar (see Fig. S1), and it also adopts a special conformation generally not allowed for non-glycine amino acids. In the present study, the roles of Arg60 and Tyr61 were investigated by site-directed mutagenesis. Among these mutants, Y61G showed significantly higher activity than did the wild type, and it was further characterized. The improved activity of *Tm*Cel12A can reduce the production cost and thus encourages its industrial use.

Materials and methods

Site-directed mutagenesis, protein expression, and purification

All chemicals used here were purchased from Sigma–Aldrich if not mentioned otherwise. The *Tm*Cel12A pET16b plasmid was constructed as described previously (Cheng et al. 2011). Each mutant was prepared by using QuickChange site-directed mutagenesis kit (Agilent) with *Tm*Cel12A pET16b as the template. The sequences of the mutated primers are listed in Supplementary Table S1. These constructs were transformed separately into *Escherichia coli* BL21 (DE3) strain, the bacteria were propagated in LB at 37°C until OD_{600nm} reached 0.6, and the mutated proteins were overexpressed at 20°C for 2 days by induction using 1 mM IPTG. All proteins were purified by FPLC system using nickel–nitriloacetic acid (Ni–NTA) and diethylaminoethyl (DEAE) columns (GE Healthcare). The purification conditions for the mutant proteins were similar to those for the wild type (Cheng et al. 2011). The buffers and gradient were 25 mM Tris (J.T. Baker), 150 mM NaCl, pH 7.5, and 20–250 mM imidazole for the Ni–NTA column and 25 mM Tris, pH 7.5, and 0–250 mM NaCl for the DEAE column. The proteins were eluted at about 75 mM imidazole and 125 mM NaCl, respectively. The purity of protein was checked using NuPAGE (Invitrogen) and the concentration determined by using Coomassie Plus "Bradford" Protein Assay (Thermo). The protein samples were finally concentrated to 5 mg/ml in 25 mM Tris, 150 mM NaCl, pH 7.5, by using Amicon Ultra-15 Centrifugal Filter Units (Millipore).

Cellulase activity assay and kinetic analysis

The assay procedure was similar to that of König et al. (2002). In general, 200 μl of the protein solution with a proper

concentration as diluted in 50 mM sodium acetate, pH 5.0, was mixed with an equal volume of 1% β-glucan substrate (Megazyme) and incubated in an 85°C water bath for 10 min. The solution was then mixed with 600 μl 1% DNS and then incubated at 100°C in boiling water for 5 min. The absorbance at 540 nm was then measured for calculation of the enzyme activity. A standard curve for calibrating the enzyme activity was prepared by using a series of 0–0.25 mg/ml glucose solutions. One unit of activity is defined as the amount of enzyme that releases 1 μmol product per minute.

For the kinetic analysis, optimal protein concentration was first determined by using a series of 60–1,200 ng/ml protein solutions and 20 mg/ml β-glucan substrate. The enzyme activity was then measured by using the optimal level (120 ng/ml) of protein and a series of 1–20 mg/ml β-glucan solutions. Based on these data, the kinetic parameters were obtained by using the Michaelis–Menten model and curve-fitting analysis with a computer.

Optimal reaction temperature and thermostability analyses

For the optimal temperature analysis, individual protein samples of 150 ng/ml concentration were mixed with equal volumes of 1% β-glucan and incubated for 10 min at different working temperatures. The enzyme activity was then determined by dinitrosalicylic acid (DNS) assay of the product as described above. For the thermostability analysis, both the wild-type and mutant proteins (150 ng/ml) were incubated at 85°C or 95°C water bath for different times and subsequently cooled on ice for 5 min and recovered at room temperature for another 5 min. The enzyme activity at 85°C was then determined by DNS assay.

Crystallization and data collection

All protein crystals were obtained by using sitting drop vapor diffusion method at room temperature, under similar conditions as for the wild-type enzyme (Cheng et al. 2011). The reservoir for growing the monoclinic Y61G crystals contained 0.1 M Bis–Tris, pH 5.5; 0.2 M ammonium sulfate (J.T. Baker); and 25% PEG3350. The crystals reached a suitable size in 2 days. Before flash-cooling to cryogenic temperatures, the crystals were soaked in a cryoprotectant solution of 0.1 M Bis–Tris, pH 5.5; 0.2 M ammonium sulfate; 5% glycerol; and 30% PEG3350. The Y61G–substrate complex crystal was obtained by soaking the Y61G crystal in the cryoprotectant solution which contained 10 mM cellotetraose.

For the orthorhombic Y61GG crystals, a slightly different reservoir that contained 0.1 M Bis–Tris, pH 5.5; 0.2 M ammonium sulfate; 5% glycerol; and 21% PEG3350 was used. The Y61GG-cellobiose crystal was prepared by soaking in the reservoir solution that contained 10 mM cellobiose. Inclusion of 10 mM cellotetraose in the protein solution of Y61GG resulted in a different monoclinic crystal form. The reservoir for growing the Y61del crystals contained 0.1 M Bis–Tris, pH 5.5; 0.15 M ammonium sulfate; 5% glycerol; and 20% PEG3350, and the protein solution contained 10 mM cellobiose. The resulting Y61del crystals were similar to the monoclinic Y61GG crystals. No cryoprotectant was necessary for the Y61GG and Y61del crystals.

The X-ray diffraction data from the crystals were collected at beam line BL13B1 of the National Synchrotron Radiation Research Center in Hsinchu, Taiwan, and at beam line BL17U of the Shanghai Synchrotron Radiation Facility in Shanghai, China. The data were processed by using HKL2000 (Otwinowski and Minor 1997).

Structure determination and refinement

All three crystal forms obtained in this study, including the monoclinic Y61G crystals, the orthorhombic Y61GG crystals, and the monoclinic crystals of Y61GG and Y61del, are different from those used in the previous study (Table 1). The structures were solved by the molecular replacement method using the program CNS (Brunger et al. 1998). Subsequent crystallographic refinement and manual adjustment of the models were carried out by using the programs CNS, O, and COOT (Brunger et al. 1998; Emsley et al. 2010; Jones et al. 1991). The structural diagrams were drawn by using PyMol (http://pymol.sourceforge.net).

Accession numbers

Coordinates and structure factors of the Y61G-cellotetraose, Y61GG, and Y61GG-cellobiose crystals have been deposited in the Protein Data Bank, with accession numbers 3VHN, 3VHO, and 3VHP.

Results

Activity of wild-type *Tm*Cel12A and the mutant enzymes

Our previous study of *Tm*Cel12A–substrate complex structures showed that Arg60 and Tyr61 participate in binding to the substrate (Cheng et al. 2011). To further investigate the roles of these two residues, each was replaced with different amino acids by site-directed mutagenesis. Arg60 was mutated to another positively charged residue in R60K, and the side chain was reduced to a methyl group in R60A. More mutants were produced for Tyr61, which was replaced by an arginine (Y61R), a tryptophan (Y61W), a phenylalanine (Y61F), an alanine (Y61A), or a glycine (Y61G). As shown in Fig. 1, the activity of each mutant deviated significantly from that of the wild type. The mutants R60K, R60A, Y61R, and Y61W showed lower activity than did the wild type, but higher

Table 1 Data collection and refinement statistics

	Y61G-cellotetraose	Y61GG (native)	Y61GG-cellobiose
Data collection			
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimension a, b, c (Å)	91.9, 124.0, 127.6	45.7, 74.3, 168.9	45.3, 74.5, 165.0
α, β, γ (°)	90.0, 107.1, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution range (Å)	25–2.5 (2.59–2.50)	25–1.93 (2.00–1.93)	25–1.93 (2.00–1.93)
Unique reflections	91,224 (7,983)	42,568 (4,259)	42,524 (4,221)
Redundancy	2.9 (2.4)	5.7 (5.9)	7.2 (7.5)
Completeness (%)	97.7 (85.9)	95.5 (96.8)	99.1 (100.0)
Average $I/\sigma(I)$	11.0 (2.4)	23.0 (8.4)	42.0 (21.4)
R_{merge} (%)	10.8 (29.3)	8.0 (21.6)	4.9 (11.6)
Refinement			
Positive reflections	85,491 (6,058)	41,307 (4,033)	41,782 (4,041)
R_{work} (%)	18.5 (31.2)	17.5 (19.1)	17.9 (17.2)
R_{free} (%)	24.5 (36.0)	21.6 (24.1)	22.0 (24.4)
RMSD bond lengths (Å)	0.016	0.017	0.016
RMSD bond angles (°)	1.9	1.8	1.8
Dihedral angles (%) in			
Preferred regions	95.1	96.9	96.3
Allowed regions	4.2	2.8	3.1
Average B (Å^2)/non-H atoms			
Protein	33.1/16,676	19.8/4,177	20.1/4,188
Water molecules	40.3/863	33.2/484	29.9/392
Carbohydrate	30.9/184		17.6/46
PDB ID code	3VHN	3VHO	3VHP

Values in parentheses are for the highest-resolution shells. See Table S2 for the statistics of the monoclinic Y61GG and Y61del crystals

activity was observed in the mutants Y61F, Y61A, and Y61G. Notably, the activity of Y61G was enhanced by 70%.

Intrigued by the fact that the peptide dihedral angles of Tyr61 were located in a disallowed region of the Ramachandran plot, indicating that Tyr61 had a special conformation not frequently observed for non-glycine amino acid residues, we further produced an insertion and a deletion mutant based on this particular mutant of Y61G. Both proteins of the insertion mutant Y61GG (substitution of Tyr61 by two glycine residues) and the deletion mutant Y61del (deletion of Tyr61) showed much lower activity than did the wild-type enzyme (Fig. 1). These results suggest that the precise length of the A3-B3 loop is critical to the activity of *Tm*Cel12A.

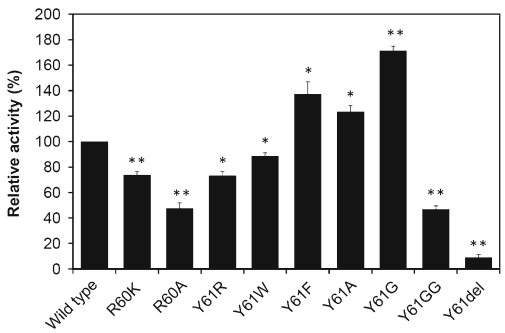

Fig. 1 Comparative activity analysis of the wild-type *Tm*Cel12A and the mutant enzymes. Each protein sample was normalized to a concentration of 150 ng/ml. The relative activity of each mutant enzyme was expressed as a percentage of the wild-type enzyme's activity. The standard error of the mean (SEM) was also shown for each mutant. All measurements were performed in triplicates. The two-tailed *P* values were determined by an unpaired Student's *t* test (*$P<0.05$; **$P<0.001$)

Kinetic analysis of the wild-type enzyme and the Y61G mutant

The catalytic properties of the wild-type and mutant enzymes were assessed by measuring the reaction velocity at different concentrations of β-glucan, which served as the substrate. The kinetic parameters calculated using these data showed that the Y61G mutant had a higher catalytic rate (k_{cat}) than did the wild-type enzyme while the K_m value of the Y61G mutant was also higher than that of the wild type (Table 2). Higher K_m of an enzyme indicates lower affinity

Table 2 Kinetic parameters of the wild-type *Tm*Cel12A and the Y61G mutant

	k_{cat} (s^{-1})	K_m (mg ml^{-1})	k_{cat}/K_m (ml s^1 mg^{-1})
Wild type	791±16	2.08±0.33	402±50
Y61G	1,155±42	3.75±0.32	313±24

to the substrate. In other words, the Y61G mutant had a lower affinity to the substrate than did the wild-type enzyme and showed lower activity at reduced substrate concentrations. Therefore, the apparent catalytic efficiency (k_{cat}/K_m) of Y61G was not better than that of the wild type. Nevertheless, with the presence of sufficient substrate, the specific activity of Y61G was much higher than that of the wild-type *Tm*Cel12A (Fig. 1). In fact, a recent study showed that the use of k_{cat}/K_m values is not always suitable to compare different enzyme mutants (Eisenthal et al. 2007).

Effects of Gly61 on optimal reaction temperature and thermostability

The enhanced endoglucanase activity of Y61G makes it a promising enzyme for industrial use, which often involves high temperatures. To investigate whether this mutant enzyme works well as does the wild type at high temperatures and also retains the original thermostability, we measured and compared the specific activity of each enzyme at different temperatures. As shown in Fig. 2, the optimal reaction temperature analysis suggested that the endoglucanase activity of Y61G was significantly higher than the wild type when measured at every temperature point from 55°C to 95°C under the same condition with the same protein and substrate concentrations. Besides, the mutant protein showed a broader range of optimal temperatures (between 90°C and 100°C) than did the wild type (nearly 100°C).

On the other hand, as shown in Fig. 3, the thermostability profile of Y61G is similar to that of the wild-type enzyme when treated at either 95°C or 85°C. The Y61G mutant and the wild type both retained 50% activity after incubation at 95°C for 40 min or at 85°C for 8 h (Fig. 3). These results suggest that the single specific mutation of Y61G not only significantly increased enzyme activity at different reaction temperatures but also remained active at various elevated temperatures. Both enzymes can function very well at a temperature above 90°C.

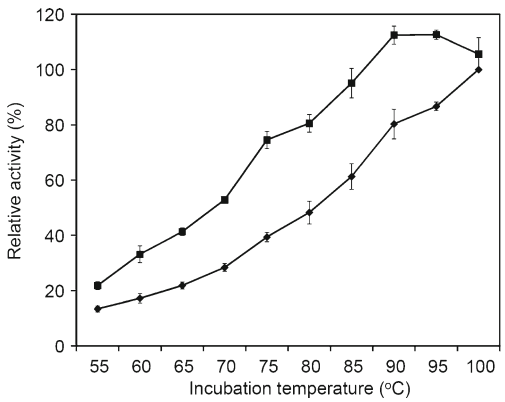

Fig. 2 Optimal reaction temperatures of the wild-type *Tm*Cel12A and the Y61G mutant. Similar to the measurement in Fig. 1, protein samples were incubated individually at different working temperatures (ranging from 55°C to 100°C in 5° steps) with the substrate. The relative activity of each enzyme at each working temperature was expressed as a percentage of the wild-type enzyme's activity at 100°C. The *line* connecting *diamonds* is for the wild-type enzyme, and that connecting *squares* is for the mutant Y61G. Each experiment was performed in triplicate, from which the SEM was calculated

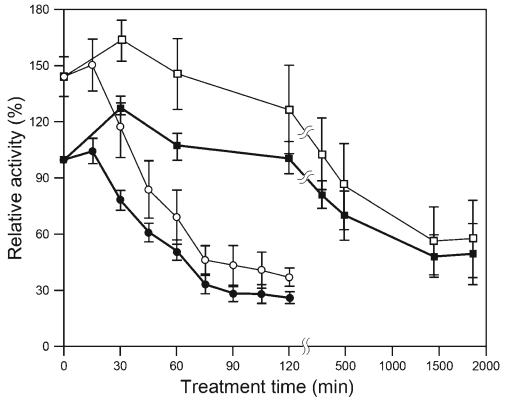

Fig. 3 Thermostability profiles of the wild-type *Tm*Cel12A and the Y61G mutant. Each sample was incubated in a water bath of either 95°C (*circles*) or 85°C (*squares*) for different time intervals of up to 120 min (2 h; at 95°C) or 1,860 min (31 h; at 85°C) and then cooled to room temperature before activity measurement at 85°C. The *bold lines* connecting *filled squares* and *circles* represent the profile of the wild-type enzyme, whereas the *thin lines* connecting *open squares* and *circles* are for the mutant. The *error bars* indicate the SEM values from triplicate experiments

Structural analyses of the Y61G, Y61GG, and Y61del crystals

To investigate if the mutations caused significant structural variation in the *Tm*Cel12A protein and its interaction with substrate, we further obtained the Y61G, Y61GG, and Y61del crystals and analyzed them by X-ray diffraction. Although similar crystallization conditions were employed, the mutant crystals turned out to have different unit-cell contents and belong to different space groups from those of the wild type. There were three different crystal forms: (1) the monoclinic crystals of Y61G, (2) the orthorhombic crystals of Y61GG, and (3) the monoclinic crystals of Y61GG and Y61del, which were different from the first Y61G crystals. Crystals of the third form were probably not single, although they diffracted X-rays to 1.5- and 1.6-Å resolution. These structures could not be refined to yield satisfactory results, but they still showed some interesting features.

The monoclinic Y61G crystals contained eight protein molecules in an asymmetric unit. The initial *R* value was 0.29 upon rigid-body refinement. Moderate non-crystallographic symmetry (NCS) restraints were employed to compensate for the mediocre resolution of 2.5 Å. There were two and four protein molecules per asymmetric unit in the orthorhombic and monoclinic

Y61GG crystals, respectively, of which the initial R values were 0.32 and 0.44. No NCS restraint was used. The Y61del crystals were isomorphous to the latter Y61GG crystals. Although the monoclinic crystals of Y61GG and Y61del were not single, the models were nevertheless refined to 2.0-Å resolution, yielding R and R_{free} of 0.33–0.35 and 0.39–0.41. Some statistics of the data collection and structural refinement can be found in Table 1 and Supplementary Table S2. See Table S3 for a summary of the resulting models.

The eight protein molecules in the Y61G crystal differed from the wild-type *Tm*Cel12A model by root-mean-square deviations (RMSD) of 0.23–0.27 Å for 255 Cα atoms, and the RMSD of Cα atoms ranged from 0.13 to 0.27 Å between the eight Y61G monomers. The four protein models of the two orthorhombic crystals of Y61GG differed from the wild type by RMSD of 0.31–0.40 Å for the 252 Cα atom pairs within a matching criterion of 2 Å in the least-square fitting procedure of the program O (Jones et al. 1991), which excluded the inserted glycine residue and the preceding Glu59 and Arg60. The four Y61GG models differed from one another by RMSD of 0.34–0.36 Å for 256 Cα atoms. The RMSD of the monoclinic Y61GG and Y61del models from the wild type were 0.28–0.31 Å for 252 Cα atoms within 2 Å, and the RMSD between the monoclinic models were 0.14–0.19 Å for 256 or 254 Cα atoms. As shown in Fig. 4a, the mutant structures were identical to that of the wild type in terms of the overall protein fold, except for the region of insertion and deletion.

The Y61G crystals were soaked with cellotetraose before data collection, but in the active site of each of the eight enzyme molecules, only a cellobiose was observed. The electron density map clearly showed that the cellobiose was bound to the −2 and −1 subsites (Fig. S2), and there was no additional sugar density in the other subsites. The interactions between the bound −2 and −1 sugar units and the enzyme, as shown in Fig. 4b, are virtually identical to those observed previously in the wild-type *Tm*Cel12A-cellotetraose crystal (Cheng et al. 2011). In the orthorhombic Y61GG crystals, the loop structure was different from that in the wild-type and Y61G crystals, but its conformation was not changed upon binding to cellobiose, which was introduced by soaking. The side chain of Arg60 deviated from its original position and did not form proper hydrogen bonds with the sugar units in the −2 and −1 subsites (Fig. 4c). The monoclinic Y61GG and Y61del crystals were obtained in the presence of cellotetraose and cellobiose, respectively. Each of the four Y61GG molecules contained a cellobiose bound to the −2 and −1 subsites, but only a glucose was observed in the −2 site of one of the four Y61del molecules (Fig. S3). The loop conformations were different in these crystals. In Y61GG-cellotetraose cocrystal, the Arg60 side chain resumed the hydrogen bonding interactions with the sugars, although in a slightly varied way, whereas the Arg60 of Y61del seemed not to be forming the proper bonds.

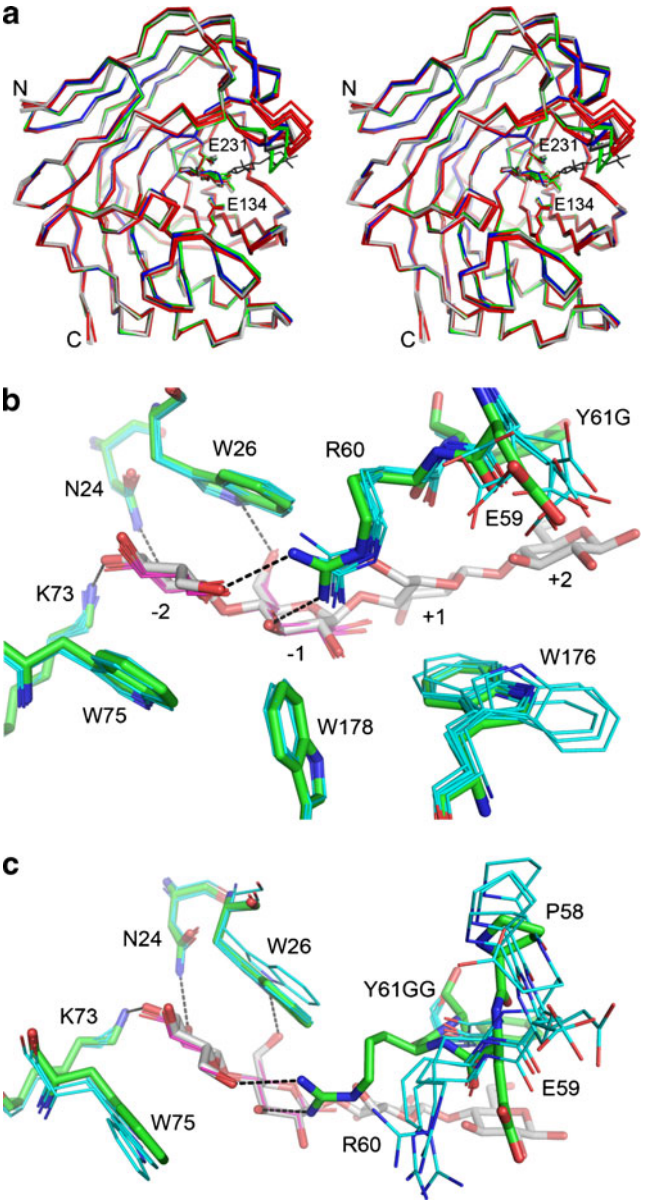

Fig. 4 Structural comparison of the wild-type and mutant *Tm*Cel12A complexes. **a** The protein molecules are drawn as Cα tracings. The Glu134 and Glu231 side chains and the bound ligands are shown as *stick models*. The wild-type structure is colored in *dark gray*; those of Y61G are in *light gray*. The orthorhombic Y61GG (apo form and cellobiose complex) structures are in *red*, those of the monoclinic Y61GG in *green*, and Y61del in *blue*. **b** The bound sugar in the wild-type complex and some surrounding amino acid residues are shown as *thick stick models* with the carbon atoms colored in *gray* and *green*. Those of Y61G are shown as *thin sticks* using *pink* and *cyan* carbon atoms. The hydrogen bonds are denoted by *dashed lines*. **c** The wild-type and the orthorhombic Y61GG models are shown in a similar manner as in (**b**)

Discussion

Enzyme–substrate interactions in the mutants

In this study, a few mutants at Arg60 and Tyr61 in a unique surface loop of *Tm*Cel12A were produced and tested for their activity. This loop is adjacent to the active site, but it does not participate directly in the catalytic reaction. That mutating Arg60 to a lysine or an alanine decreased the enzyme's activity by about 30% or 50% is possibly due to weaker enzyme–substrate interactions. In the wild-type enzyme, Arg60 not only makes four direct hydrogen bonds with three sugar units (Fig. S1) but also forms an enclosure by contacting Trp178 on the opposite side of the substrate-binding cleft (Cheng et al. 2011). The side chain of Lys60 in the mutant R60K, although as flexible and positively charged, can make fewer hydrogen bonds because it contains only one amino group. The imperfect orientation of the hydrogen bonds should also have reduced the enzyme's affinity to the substrate. We are not sure whether Lys60 forms an enclosure with Trp178 as does Arg60, but the truncated side chain of Ala60 in R60A definitely cannot make it. Neither is it likely for the three hydrogen bonds involving the Arg60 side chain to form in the R60A mutant. Thus, probably because of the weaker substrate interactions, both mutants showed lower activities.

Mutating Tyr61 into an arginine or a tryptophan decreased the activity by 30% or 10%. Both of the flat guanidinium and indole groups in the mutated side chains of Arg61 and Trp61 are capable of stacking with the substrate. However, the former side chain is much more flexible, and the latter is significantly larger than that of Tyr61. Consequently, the mutants Y61R and Y61W might bind to the substrate in slightly different orientations, disrupting the original interactions and thus affecting the ongoing catalytic reaction. On the contrary, mutation of Tyr61 to phenylalanine, alanine, or glycine improved the enzyme's activity by 40%, 20%, or 70%, respectively. The removal of the hydroxyl group in the mutant Y61F seems not to perturb the stacking interactions at all, but it may lose some secondary hydrogen bonds with the substrate (beyond the +2 sugar). The methyl group of Ala61 in Y61A should remain in contact with the substrate. The lack of stacking interactions may reduce the substrate affinity, but it may also facilitate release of the reducing-end sugars after cleavage. Interestingly, the best activity was observed in the mutant Y61G. It might be attributed in part to the fact that Tyr61 actually has a glycine-like conformation in the wild-type enzyme. Mutating it to Gly61 may stabilize the loop conformation and its interactions with the substrate. In summary, the effect of these mutants on the activity of *Tm*Cel12A is a subtle balance between the affinity for substrate binding and the readiness for product release.

Structural basis of the enhanced activity of Y61G

The bound cellobiose to the −2 and −1 subsites of each enzyme molecule in the Y61G-cellotetraose crystal was probably a result of hydrolysis. In the wild-type *Tm*Cel12A–cellotetraose complex crystal, which was also obtained by soaking, only partial hydrolysis occurred, and sugar units bound to the +1 and +2 subsites were observed (Cheng et al. 2011). Probably cellotetraose was still bound to the four subsites from −2 to +2, but the cleaved moiety at the reducing end was released from the active site of Y61G because the stabilizing interactions were weaker than those in the wild-type enzyme. The various side-chain conformations of the nearby Glu59 and Trp176 (Fig. 4b) are another evidence of the empty +1 and +2 subsites and also imply the increased entropy upon product release. These observations are consistent with the increased k_{cat} and K_m values. For industrial applications, the reaction conditions often involve high temperature and high substrate concentration, which compensate for the higher K_m value. We have shown that the thermostability profiles of the wild-type and Y61G enzymes are similar, and the optimal temperature of Y61G was broadened by 10°. Probably, this mutant enzyme is more useful than the wild type. Indeed, the specific activity of Y61G was significantly higher (up to twofold) than the wild type when expressed in *Pichia pastoris* by using a 50-L fermentor (data not shown), and similar results were obtained when the expression was in *E. coli*.

The insertion mutant Y61GG showed a 50% reduction in activity. In the orthorhombic crystal, the adjacent residues in the loop moved away from the substrate (Fig. 4c), but in the monoclinic crystal, they were rearranged to form a set of new hydrogen bonds between Arg60 and the substrate (Fig. S3A). Although the inserted glycine residue locally perturbed the original loop conformation, thanks to the increased length and presumably the increased flexibility of the loop, the mutant enzyme was able to reestablish some alternative interactions with the substrate. The rearrangement nonetheless would need to overcome some energy barrier, and thus, the Y61GG recovered one half of the original activity, comparable to that of R60A, in which all those hydrogen bonds were absent. By contrast, the shorter loop of the deletion mutant Y61del lacked such flexibility. The structure of its cocrystal with cellobiose showed that the side chain of Arg60 tended to occupy the +1 subsite, and it would interfere with real substrate binding (Fig. S3B). Consequently, the mutant enzyme retained only 10% activity.

Comparison with other studies of GH11 and GH12 enzymes

In contrast to the enhancement of enzymatic activity presented here, most of the previous studies about structural-based modification of the GH12 enzymes were to improve thermostability, especially using *T. reesei* Cel12A. Many industrial cellulases

are obtained from *T. reesei* because of the high production level and enzyme activity. Because *Tr*Cel12A shows low thermostability, its structure was compared with those of the *Streptomyces* and *Humicola grisea* enzymes, and the thermostability was improved by site-specific mutations (Sandgren et al. 2003a, b). In a more recent study, an engineered mesoactive-thermostable protein was designed by combining the groove region from the mesophilic *Tr*Cel12A with the complementary regions from the thermophilic *Rhodothermus marinus* Cel12A (Kapoor et al. 2008).

On the other hand, the roles of aromatic residues in the sugar-binding subsites of *Bacillus subtilis* XynA, which is a GH11 enzyme that belongs to the same clan-C glycoside hydrolases and folds into a β-jellyroll architecture as does Cel12A, were investigated by site-directed mutagenesis (Pollet et al. 2010). The results indicated that mutating aromatic residues to more neutral residues not necessarily enhances activity, but can also decrease activity. The observed differences might be explained by the different types of interactions between the aromatic residues and sugars. GH11 has also seen much more structure-based modifications for its thermostability and pH resistance than for the enzymatic activity. Mutating a glycine residue to cysteine in the interior hydrophobic region (You et al. 2010) or introduction of arginines by substituting serine and threonine residues on the surface (Sriprang et al. 2006) improved the enzyme's thermostability. Comparison of the sequences and structures between the GH11 xylanases with different pH optimum has also allowed the identification of specific residues which is important to the pH stability (Al Balaa et al. 2009; De Lemos Esteves et al. 2005).

In conclusion, despite the fact that several protein structures of GH12 family cellulases have been solved in recent years, further structure-based modifications are still few. Most of those studies were focused on thermostability rather than enzymatic activity. Here, the *Tm*Cel12A mutant Y61G in the unique surface loop of A3-B3, with an enhanced activity, represents a successful example of protein engineering, and the conserved hyperthermostability makes it suitable for industrial application. Although the mutant has a slightly reduced affinity, it can be compensated by high substrate concentration. By contrast, the wild-type enzyme should have been optimized for the lower substrate concentration usually found in natural environment, which can be very different than the highly reactive conditions employed in the industry.

Acknowledgments We thank the National Synchrotron Radiation Research Center of Taiwan and the Shanghai Synchrotron Radiation Facility of China for beam-time allocations and data-collection assistance. This work was supported by grants from the National Science Council of Taiwan (NSC 98-2313-B002-033-MY3 to JRL), the National Basic Research Program of China (2011CB710800 to RTG), and the Tianjin Municipal Science and Technology Commission (10ZCKFSY06000 to RTG).

References

Al Balaa B, Brijs K, Gebruers K, Vandenhaute J, Wouters J, Housen I (2009) Xylanase XYL1p from *Scytalidium acidophilum*: site-directed mutagenesis and acidophilic adaptation. Bioresour Technol 100:6465–6471

Bhat MK (2000) Cellulases and related enzymes in biotechnology. Biotechnol Adv 18:355–383

Birsan C, Johnson P, Joshi M, MacLeod A, McIntosh L, Monem V, Nitz M, Rose DR, Tull D, Wakarchuck WW, Wang Q, Warren RA, White A, Withers SG (1998) Mechanisms of cellulases and xylanases. Biochem Soc Trans 26:156–160

Blumer-Schuette SE, Kataeva I, Westpheling J, Adams MW, Kelly RM (2008) Extremely thermophilic microorganisms for biomass conversion: status and prospects. Curr Opin Biotechnol 19:10–217

Bornscheuer UT, Pohl M (2001) Improved biocatalysts by directed evolution and rational protein design. Curr Opin Chem Biol 5:137–143

Böttcher D, Bornscheuer UT (2010) Protein engineering of microbial enzymes. Curr Opin Microbiol 13:274–282

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D 54:905–921

Cheng YS, Ko TP, Wu TH, Ma Y, Huang CH, Lai HL, Wang AHJ, Liu JR, Guo RT (2011) Crystal structure and substrate-binding mode of cellulase 12A from *Thermotoga maritima*. Proteins 79:1193–1204

Crennell SJ, Cook D, Minns A, Svergun D, Andersen RL, Nordberg Karlsson E (2006) Dimerisation and an increase in active site aromatic groups as adaptations to high temperatures: X-ray solution scattering and substrate-bound crystal structures of *Rhodothermus marinus* endoglucanase Cel12A. J Mol Biol 356:57–71

De Lemos Esteves EF, Gouders T, Lamotte-Brasseur J, Rigali S, Frere JM (2005) Improving the alkalophilic performances of the Xyl1 xylanase from *Streptomyces* sp. S38: structural comparison and mutational analysis. Protein Sci 14:292–302

Dumon C, Varvak A, Wall MA, Flint JE, Lewis RJ, Lakey JH, Morland C, Luginbuhl P, Healey S, Todaro T, DeSantis G, Sun M, Parra-Gessert L, Tan X, Weiner DP, Gilbert HJ (2008) Engineering hyperthermostability into a GH11 xylanase is mediated by subtle changes to protein structure. J Biol Chem 283:22557–22564

Dutta T, Sahoo R, Sengupta R, Ray SS, Bhattacharjee A, Ghosh S (2008) Novel cellulases from an extremophilic filamentous fungi *Penicillium citrinum*: production and characterization. J Ind Microbiol Biotechnol 35:275–282

Eisenthal R, Danson MJ, Hough DW (2007) Catalytic efficiency and kcat/Km: a useful comparator? Trends Biotechnol 25:247–249

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D 66:486–501

Henrissat B, Teeri TT, Warren RA (1998) A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. FEBS Lett 425:352–354

Hua C, Yan Q, Jiang Z, Li Y, Katrolia P (2010) High-level expression of a specific β-1,3-1,4-glucanase from the thermophilic fungus *Paecilomyces thermophila* in *Pichia pastoris*. Appl Microbiol Biotechnol 88:509–518

Ilari A, Fiorillo A, Angelaccio S, Florio R, Chiaraluce R, van der Oost J, Consalvi V (2009) Crystal structure of a family 16 endoglucanase from the hyperthermophile *Pyrococcus furiosus*: structural basis of substrate recognition. FEBS J 276:1048–1058

Jarvis M (2003) Chemistry: cellulose stacks up. Nature 426:611–612

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 47:110–119

Kapoor D, Kumar V, Chandrayan SK, Ahmed S, Sharma S, Datt M, Singh B, Karthikeyan S, Guptasarma P (2008) Replacement of the active surface of a thermophile protein by that of a homologous mesophile protein through structure-guided 'protein surface grafting'. Biochim Biophys Acta 1784:1771–1776

König J, Grasser R, Pikor H, Vogel K (2002) Determination of xylanase, β-glucanase, and cellulase activity. Anal Bioanal Chem 374:80–87

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577

Ohmiya K, Sakka K, Karita S, Kimura T (1997) Structure of cellulases and their applications. Biotechnol Genet Eng Rev 14:365–414

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276:307–326

Percival-Zhang YH, Himmel ME, Mielenz JR (2006) Outlook for cellulase improvement: screening and selection strategies. Biotechnol Adv 24:452–481

Pollet A, Lagaert S, Eneyskaya E, Kulminskaya A, Delcour JA, Courtin CM (2010) Mutagenesis and subsite mapping underpin the importance for substrate specificity of the aglycon subsites of glycoside hydrolase family 11 xylanases. Biochim Biophys Acta 1804:977–985

Sandgren M, Gualfetti PJ, Paech C, Paech S, Shaw A, Gross LS, Saldajeno M, Berglund GI, Jones TA, Mitchinson C (2003a) The *Humicola grisea* Cel12A enzyme structure at 1.2 A resolution and the impact of its free cysteine residues on thermal stability. Protein Sci 12:2782–2793

Sandgren M, Gualfetti PJ, Shaw A, Gross LS, Saldajeno M, Day AG, Jones TA, Mitchinson C (2003b) Comparison of family 12 glycoside hydrolases and recruited substitutions important for thermal stability. Protein Sci 12:848–860

Schülein M (2000) Protein engineering of cellulases. Biochim Biophys Acta 1543:239–252

Selinger LB, Forsberg CW, Cheng KJ (1996) The rumen: a unique source of enzymes for enhancing livestock production. Anaerobe 2:263–284

Sriprang R, Asano K, Gobsuk J, Tanapongpipat S, Champreda V, Eurwilaichitr L (2006) Improvement of thermostability of fungal xylanase by using site-directed mutagenesis. J Biotechnol 126:454–462

Szijártó N, Horan E, Zhang J, Puranen T, Siika-Aho M, Viikari L (2011) Thermostable endoglucanases in the liquefaction of hydrothermally pretreated wheat straw. Biotechnol Biofuels 4:2

Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1–43

Viikari L, Alapuranen M, Puranen T, Vehmaanpera J, Siika-Aho M (2007) Thermostable enzymes in lignocellulose hydrolysis. Adv Biochem Eng Biotechnol 108:121–145

Vocadlo DJ, Davies GJ (2008) Mechanistic insights into glycosidase chemistry. Curr Opin Chem Biol 12:539–555

Yang S, Wang Y, Jiang Z, Hua C (2008) Crystallization and preliminary X-ray analysis of a 1,3-1,4-β-glucanase from *Paecilomyces thermophila*. Acta Crystallogr F 64:754–756

Yang D, Weng H, Wang M, Xu W, Li Y, Yang H (2010) Cloning and expression of a novel thermostable cellulase from newly isolated *Bacillus subtilis* strain I15. Mol Biol Rep 37:1923–1929

You C, Huang Q, Xue H, Xu Y, Lu H (2010) Potential hydrophobic interaction between two cysteines in interior hydrophobic region improves thermostability of a family 11 xylanase from *Neocallimastix patriciarum*. Biotechnol Bioeng 105:861–870