Journal of Biotechnology 306 (2019) 118–124

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Improvement in activity of cellulase Cel12A of *Thermotoga neapolitana* by error prone PCR

Abdul Basit[a,c,d], Razia Tajwar[a], Saima Sadaf[b], Yang Zhang[c], M. Waheed Akhtar[a,*]

[a] School of Biological Sciences, University of the Punjab, Lahore, 54590, Pakistan
[b] Institute of Biochemistry and Biotechnology, University of the Punjab, Lahore, Pakistan
[c] Department of Computational Medicine & Bioinformatics, University of Michigan, MI, United States
[d] Department of Biotechnology, Virtual University of Pakistan, 1-Davis Road, Lahore, Pakistan

ARTICLE INFO

Keywords:
Error prone PCR
Cellulase
Cel12A
Thermotoga neapolitana

ABSTRACT

Using multi-step error prone PCR (ep-PCR) of the gene encoding endoglucanase Cel12A (27 kDa) from *Thermotoga neapolitana,* mutants were obtained with many fold increase in the enzyme activity. The best mutant (C6, N47S/E57 K/ V88A/S157 P/K165 H) obtained after four rounds of ep-PCR showed 2.7 −, 5 − and 4.8 − fold increase in activity against CMC, RAC and Avicel, respectively, compared with the wild type enzyme. The other characteristics of the mutated enzyme with respect to stability, optimum working pH and temperature were comparable to the wild type enzyme.C6 mutant showed higher binding efficiency towards the rice straw (~50%) than the wild type (~41%). The structural information obtained from the protein docking of the wild type Cel12A and its mutant showed that E57 K improved the binding affinity between enzyme and ligand by producing conformational changes in the catalytic cavity. The other mutations can facilitate the enzyme-substrate binding interactions to enhance catalytic activity although they are not directly involved in catalysis. The wild type and mutant enzyme produce cellobiose as the major products for both soluble and insoluble substrates, suggesting that this enzyme should be a cellobiohydrolase instead of endoglucanase as previously reported.

1. Introduction

Plant biomass, a rich source of cellulose and hemicellulose, is the most abundant renewable resource found on earth (Wang et al., 2005). The interest in the bioconversion of plant biomass has increased in the recent past. Three types of enzyme activities are required for breakdown of cellulose into fermentable sugar: endoglucanase cleaves the cellulose internally by breaking β 1–4 glycosidic bond, exoglucanase cleaves cellulose at their ends releasing cellobiose and β-glucosidase cleaves cellobiose into glucose which can be fermented into ethanol (Liang et al., 2011). Due to the low catalytic efficiency of the enzymes, hydrolysis of plant biomass polysaccharides has been an expensive processing step in biofuel production (Wang et al., 2010). Application of protein engineering methods is a highly potential approach for improving the enzyme activities. Optimizing the hydrolysis of cellulose substrate by improving the activity of cellulases has remained a major focus of the research (Liu et al., 2011).

Directed evolution is a quick and cost effective method of producing mutant enzymes with wide range of diversity (Rubin-Pitel et al., 2006). Directed evolution has been widely used in the recent past to engineer the cellulolytic enzymes for improving activity using ep-PCR (Pritchard et al., 2005; Ahmad et al., 2014). A library of mutants generated by ep-PCR is screened on solid agar plate containing CMC followed by staining with Congo red dye to select the mutants with improved activity (Farrow and Arnold, 2011).

Previously random mutagenesis of exoglucanase, Cel9 from *Clostridium phytofermentans* showed 3-fold enhancement in activity against microcrystalline cellulose substrate, compared with the wild type enzyme. Most of the mutations were found in the binding domain facilitating the binding towards the substrate to improve cellulolytic activity (Ahmad et al., 2014). Mutagenesis of Cel5A of *Thermotoga maritima* showed up to 30% increase in activity on ionic-liquid pretreated switch grass and 25–42% increase on CMC. Interestingly, most of the mutations found in improved mutants were distally located from the active sites on the enzyme surface and not directly involved with substrate binding (Croft et al., 2013).

In this study Cel12A (EC 3.2.1.4; family GH12) of *T. neapolitana* and its mutants produced through random mutagenesis were expressed in *E.*

* Corresponding author.
E-mail addresses: basitbch@gmail.com (A. Basit), razia.sbs@pu.edu.pk (R. Tajwar), saima.ibb@pu.edu.pk (S. Sadaf), zhng@umich.edu (Y. Zhang), mwa.sbs@pu.edu.pk (M.W. Akhtar).

https://doi.org/10.1016/j.jbiotec.2019.09.011
Received 27 March 2019; Received in revised form 17 September 2019; Accepted 18 September 2019
Available online 21 September 2019
0168-1656/

coli and characterized.

2. Material and methods

2.1. Bacterial strains, plasmid and chemical

E. coli DH5-α was used for cloning and recombinant plasmid propagation while *E. coli* BL21 CodonPlus (RIL) for expression of the recombinant proteins. The transformed cells were grown in LB medium supplemented with 100 µg/mL ampicillin. IPTG at a concentration of 0.5 mM was used as inducer. pET-22b (+) (Novagen, Madison, USA) was used as the expression vector. Chromosomal DNA of *T. neapolitana* (DSM-3109, from DSMZ, Germany) was the source for cellulase gene. Carboxymethyl-cellulose (CMC), microcrystalline cellulose and Congo red dye were purchased from Sigma-Aldrich.

2.2. PCR amplification of Cel12A gene

The Cel12A gene was amplified from *T. neapolitana* genomic DNA trough conventional PCR using CelnA-F [CATATGGTTGAACTGACCGC ACC] as forward and CelnA-R [CTCGAGTCATCCTTTCACCTCCATATC] as reverse primers, containing *Nde*I and *Xho*I restriction sites, respectively. The primers were designed on the basis of published genes sequence (GenBank: U93354.1). PCR reaction was performed for 30 cycles with denaturation at 95 °C for 35 s, annealing at 60 °C for 35 s and extension at 72 °C for 40 s. Final extension was done at 72 °C for 15 min.

2.3. Random mutagenesis

To generate random mutagenesis library, error prone PCR was performed using DE-F [3´-GGATCCCAGCATATGGTAGAACTGACCGCA CCG-5´] as forward and DE-R [3´-GGATCCCTCGAGTCATTCTTTCACC TCCATATCG-5´] as reverse primers using genomic DNA of *T. neapolitana* as template. Overhangs of 9 and 6 base pairs were introduced to DE-F and DE-R primers, respectively, for restriction of the PCR amplified product using *Nde*I and *Xho*I restriction enzymes. Ep-PCR reaction mixture containing 4–7 mM $MgCl_2$, 0.2 mM $MnCl_2$, 5 units of *Taq* polymerase, 0.2 mM dATPs and dGTPs while 1 mM dCTPs and dTTPs were run for 30 cycles using the same PCR conditions as used for the conventional PCR. The amplified product obtained from the conventional PCR and ep-PCR, were purified and ligated in pET-22b (+) vector followed by transformation of *E. coli* BL21 CodonPlus. Transformants were grown on LB agar plate containing 100 µg mL^{-1}ampicillin at 37 °C. The colonies thus appeared were plated in duplicate on agar plate containing 0.1mM IPTG and 100 µg mL^{-1} ampicillin. One set of plates was overlaid with 0.5% agar solution containing 0.2% CMC (Sigma Inc., St. Louis) and incubated at 60 °C for 2 h, followed by staining with 0.1% Congo red solution for 15 min at room temperature. The plates were washed with 1 M NaCl. The mutants creating larger zone of hydrolysis were grown in LB broth and supernatants from their cell lysates were assayed for CMCase activity.

For efficient molecular diversity, the mutants with higher activity obtained from the first round of ep-PCR were used as template for the second round of ep-PCR according to the protocol described above. Similarly, four iterative rounds of random mutagenesis were performed followed by screening for the activity each time.

2.4. Expression and purification

Inoculation of *E. coli* cells carrying the wild type Cel12A and its selected mutants, made in LB medium was used to inoculate 400 mL LB medium containing 100 µg mL^{-1} ampicillin and incubated in orbital shaker at 37 °C. The culture at an OD_{600} ~ 1.0 was induced with 0.4 mM IPTG and incubated for another 6 h till the OD_{600} reached around 7.0. The cells were harvested by centrifugation at 9000 *g* for 20 min and resuspended in 0.05 M Tris buffer followed by ultra sonication using UP400S ultraschall processor. For partial purification, the supernatant was subjected to heat treatment at 90 °C for 30 min to denature the heat labile host cell proteins. The supernatant obtained after centrifugation was analyzed by SDS-PAGE. Percentage of the enzyme fraction in the total protein was determined by densitometric analysis using Gel Documentation System (Syngene, UK). The percentage of the enzyme in each of the sample was also determined by using GelQuan (Rehbein and Harald, 2015) and Image J software (Schneider et al., 2012).

2.5. Enzyme assays

The enzyme activities of the wild type Cel12A and its mutants were determined by mixing 0.5 mL of the enzyme sample with 0.5 mL of 1% CMC dissolved in 0.05 M Tris buffer (pH 6.0). Each of the sample was diluted to adjust the concentration of the enzyme at 80 µg/0.5 mL. The mixture was incubated for 10 min at 80 °C in a shaking water bath. The amount of the reducing sugar released was determined by the DNS reagent method (Xiao et al., 2005) using glucose (Sigma Aldrich) as the standard. A unit of the enzyme activity was defined as the amount of enzyme required to release 1 µM of reducing sugar per minute.

The activities of Cel12A and its mutants were also determined against microcrystalline cellulose (Avicel) and the regenerated amorphous cellulose (RAC). The RAC was produced by the method described previously (Ahmad et al., 2014). Reaction mixture consisted of 0.5 mL of enzyme sample and 10 mg of Avicel or RAC contained in 0.5 mL Tris buffer (pH 6.0).

2.6. Characterization of Cel12A mutant

To determine the optimum pH for activity of the mutant, the assays were done at different pH using the buffers 50 mM acetate (pH 3.0–5.0), 50 mM phosphate (pH 5.5–6.5) and 50 mM Tris-Cl (pH 7.0–8.5). To determine the effect of temperature on cellulase activity of the mutant, the assays were performed at temperatures range 50–90 °C.

Stability was determined by incubating the enzyme solution at different temperatures ranging from 50 to 95 °C for up to 120 min. The residual activity was determined against CMC after every 30 min of incubation.

K_m and V_{max} values of the mutant were determined using different CMC concentrations ranging from 2 to 20 mg mL^{-1} at the optimum temperature and pH. Lineweaver-Burk plot was used to calculate the K_m and V_{max} values. The binding efficiency of the wild type and C6 mutant was determined by mixing 50 mg of the pretreated rice straw with 200 µg mL^{-1}ofthe enzyme contained in 5 mL of 50 mM Tris buffer (pH 6.0) and the mixture was incubated at 4 °C for 6 h with gentle shaking as described before (Khan et al., 2013). The mixture was then centrifuged at 9000 *g* for 15 min and the unbound activity as well as protein concentration in the supernatant was assayed (Basit and Akhtar, 2018). All the enzymatic reactions were performed in triplicate and the mean of the values with the standard deviations less than 5% were reported.

2.7. Hydrolysate analysis by HPLC

The hydrolysis products of CMC, RAC and Avicel obtained after C6 mutant treatment were analyzed by HPLC using HPX-42A column (300 x 78 mm; Bio-Rad Laboratories, Inc., Hercules, CA). Hydrolyses of CMC, RAC and Avicel was carried out for 10 min at 80 °C in 1 mL reaction containing 0.08 mg mL^{-1} enzyme. The reaction mixtures were then centrifuged at 12,000 rpm for 1 min and the supernatants obtained were syringe filtered followed by HPLC analysis as described previously (Basit and Akhtar, 2018). The reference sample contained cellotetraose, cellotriose (Megazyme), cellobiose and glucose (Sigma-Aldrich) as the standards.

2.8. Structure analysis

The structures of the wild type Cel12A and the mutant with all five mutation sites (N47S, E57 K, V88A, S157 P and K165 H) were produced using I-TASSER, which constructs structural models by assembling continuous fragments excised from multiple threading templates identified from the PDB through replica-exchange Monte Carlo (REMC) simulations (Yang et al., 2015). Starting from the I-TASSER models, binding residues of both the wild type and mutant structures were investigated by using COACH binding site prediction tool. COACH generates all possible ligand binding sites using ligand binding templates recognized by TM-SITE and S-SITE methods on the basis of specific substructure and sequence-profile alignments (Yang et al., 2013).

In order to determine the binding affinity of both the wild type Cel12A and its mutant with a selected ligand obtained from the complex structure of Cel5A (PDB ID: 6GL0), EDock, a blind protein-ligand docking approach (available at the link https://zhanglab.ccmb.med.umich.edu/EDock/) was used. The method approximates the predicted ligand binding residues as a grid point and generates initial binding pockets from the COACH binding site predictions, followed by geometrical shape and chemical feature guided refinements. Next, REMC simulation is performed under a physical force field containing electrostatic and van der Waals forces to create more than 10,000 ligand-protein docking conformations. The conformation with the maximum X-Score was selected in this study. All the generated structures were visualized through PyMOL (Delano, 2002).

3. Results and discussion

There are several methods of directed evolution used to enhance cellulase properties. One of the most successful method is ep-PCR, which is commonly used for creation of random mutations (Kaur and Sharma, 2006). We attempted in this study random mutagenesis of Cel12A gene through ep-PCR followed by expression of the mutants in *E. coli* and screening for improved activity and other properties of the enzyme.

3.1. Mutant library construction

Cel12A is a 29 kDa thermostable cellulase produced by *T. neapolitana*. It has been reported to be stable at 106 °C for more than two hours (Bok et al., 1998). Gene fragment (0.77 kb) encoding Cel12A was amplified from the genomic DNA of *T. neapolitana* and cloned in the plasmid pET-22b (+). A library of randomly mutated genes produced after several iterative rounds of ep-PCR was used to transform *E. coli* BL21 Codon Plus. Although, unbalanced nucleotide concentrations as well as $MnCl_2$ (Cadwell and Joyce, 1992) modulate the mutation frequency in ep-PCR, iterative ep-PCR provide broader chances of mutations (Shi et al., 2014a). Over 2000 colonies screened on CMC plate, four mutants C3, C22, C28 and C6 were selected from the first, 2nd, 3rd and 4th rounds of ep-PCR, respectively (Fig. 1). These mutants were selected on the basis of larger haloes around the colony as compared to the wild type colony. Sequencing of the selected mutants showed the presence of N47S, E57 K, V88A, S157 P and K165H mutations (Fig.1, Table 1). Substitutions in each round were transferred to the mutant in the subsequent round.

3.2. Expression and activities of the mutants

The mutants C3, C22, C28 and C6, which showed higher activity than the wild type Cel12A, as found after the initial screening were expressed in *E. coli* BL21 Codon Plus. Cells after cultivation, induction, harvesting and suspending in the Tris buffer (pH 6.0), were lysed by ultrasonication. The lysate supernatant after heating at 90 °C for 1 h was analyzed by SDS-PAGE. The enzyme in lysate of C6 and C28 were found about 60% purified while those of C3, C22 and the wild type were about 50% purified, as shown in Fig. 2. The percentages of the enzyme in each sample when analyzed by GelQuan and Image J software showed quite similar results.

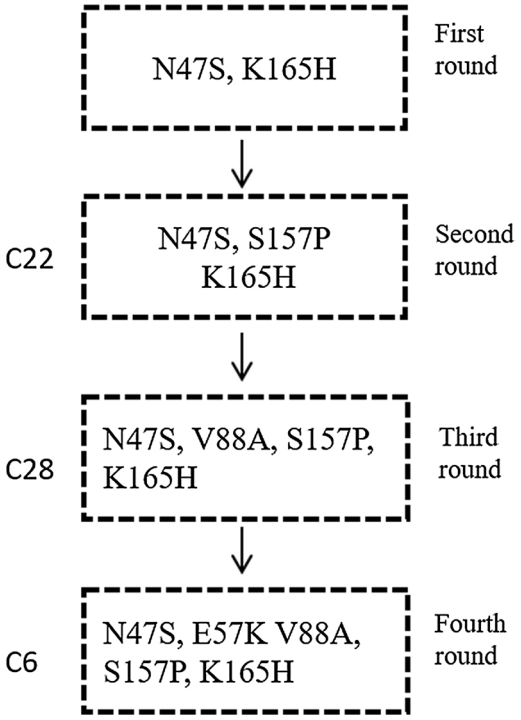

Fig. 1. Mutants showing alterations of the amino acid residues obtained after different rounds of ep-PCR.

The wild type enzyme showed specific activity 42U/mg against CMC, while C3, C22, C28 and C6 showed specific activity 48, 59, 75 and 115 U mg^{-1}, respectively (Table 1). The mutant C6 showed five mutations, which are N47S, E57 K, V88A, S157 P and K165H. Four of these *i.e.* N47S, V88A, S157 P and K165H were inherited from the previous rounds of ep-PCR. However, the E57 K substitution obtained in the last round resulted in significant increase in activity. The final round mutant C6 showed an activity 2.7 times higher than that of the wild type. This increased in activity is significantly higher as compared to 1.4-fold increase in activity reported for Cel8M of *E. coli* (Ling et al., 2016).

Against RAC and Avicel the mutant C6 showed activities of 5.0 and 4.75 U mg^{-1}, respectively, as compared to 2.31 and 2.13 U mg^{-1} of the wild type (Table 1). These activity values are higher than the previously reported activity against insoluble substrates (0.11 U mg^{-1}) of the purified Cel12A from *T. neapolitana* (Bok et al., 1998) and 30% improved activity in the case of Cel5A of *T. maritima* (Chen et al., 2013).

Generally, the activities of cellobiohydrolases reported previously were much lower as compared to those of endoglucanases. The activities of CpCel48 of *C. phytofermentans* (Zhang et al., 2010) and CBH from *Thermoascus aurantiacus* (Hong et al., 2003) were reported as 0.1875 and 0.013 U/mg, respectively, whereas C6 mutant showed much higher activity (4.75 U/mg) than all of these cellobiohydrolases. As it is known that generally the cellulases lacking a carbohydrate binding module (CBM) have lower activity against the insoluble substrates (Wang et al., 2010), fusion of a compatible CBM with the mutant C6 may enhance its activity further against microcrystalline cellulose substrates as reported for other cellulases and xylanases (Khan et al., 2013; Sajjad et al., 2010).

The activity yields of the wild type enzyme and the mutants C3, C22, C28 and C6 were 37, 39, 45, 82 and 125 units per liter per OD_{600nm} of the culture, respectively. Thus the overall activity yield of the mutant C6 was 3.37-fold higher as compared to that of the wild type

Table 1
Activity yields and specific activity of the wild type Cel12A and its mutated variants against CMC, RAC and Avicel.

Enzyme variants	Substitutions	UL^{-1} $OD_{600}{}^{-1}$ of culture	Specific activity (U/mg)		
			CMC	RAC	Avicel
Wild type	–	37 ± 2.51	42 ± 0.76	2.31 ± 0.02	2.13 ± 0.01
C3	N47S, H165K	39 ± 1.52	48 ± 1.04	2.30 ± 0.01	2.0 ± 0.05
C22	N47S, S157 P, H165K	45 ± 1.31	59 ± 1.73	3.7 ± 0.10	2.31 ± 0.01
C28	N47S, V88A, S157 P and H165K	82 ± 2.6	75 ± 1.33	4.48 ± 0.01	4.2 ± 0.09
C6	N47S, E57 K, V88A S157 P, H165K	125 ± 2.08	115 ± 4.04	5 ± 0.13	4.75 ± 0.12

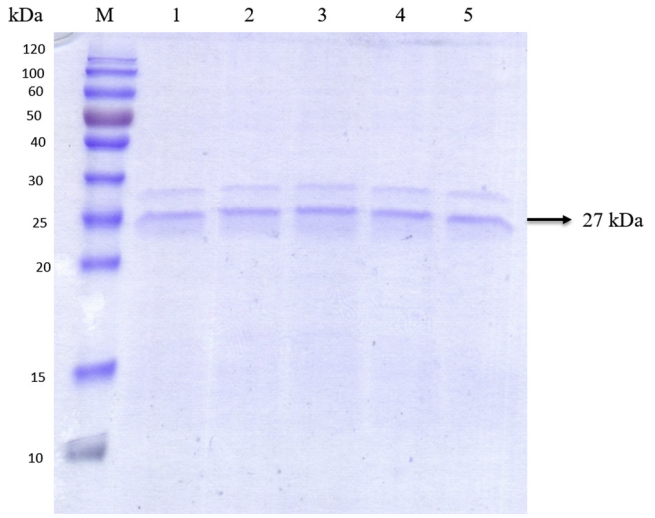

Fig. 2. SDS PAGE of the partially purified Cel12A after expression in *E. coli*. Lane M: Protein makers (Thermo Scientific™ 26,614); Lane1: C6; Lane 2: C28; Lane 3: Cel12A wild type; Lane 4: C22 and Lane 5: C3.

enzyme. Previously, directed evolution of Cel12A of *T. reesei* was reported to improve its soluble expression resulting 40-fold increases in the enzyme yield (Nakazawa et al., 2009).

3.3. Characterization of the mutant

C6 mutant was optimally active at 80 °C and showed broad pH range of 4.0–8.0 for the activity, with the maximum activity at pH 6.0 (Fig. 3a and b). These results are similar to those of the wild type Cel12A, whose optimum pH and temperature were reported previously to be 6.0 and 90 °C (Bok et al., 1998). The mutant retained more than 95% of its activity in temperature range 50–90 °C for more than 2 h, but losing 60% of the activity in incubation at 95 °C for 2 h (Fig. 3c). Another previously reported study showed that the mutant Cel16 of *Thermotoga* sp. retained more than 90% activity at 90 °C, but rapidly lost its activity at 100 °C (Shi et al., 2014a).

The K_m value for C6 mutant was found to be 28 mg mL^{-1},which is lower than the K_m value (39 mg mL^{-1}) of the wild type Cel12A. However, the V_{max} value of the wild type and C6 mutant was found to be 52 and 833 U mg^{-1}, respectively (Fig. 4), which is lower than the V_{max} value (1219 U mg^{-1}) of the wild type Cel12A purified from *T. neapolitana* (Bok et al., 1998) but significantly greater than that of the previously reported thermostable cellulases (Basit and Akhtar, 2018; Dave et al., 2015). A lower K_m and a higher V_{max} correspond with the higher activity of the C6 mutant of Cel12A.

Assay of the unbound enzyme left in the solution incubated with the pretreated rice straw showed that 40.6% and 49.8% of the total activities of the wild type enzyme and the mutant were bound to the substrate (Table 2). Higher binding affinity of C6 mutant towards the substrate corresponds to the improved activities as compared to the wild type. These results are similar to a previous study on xylanase

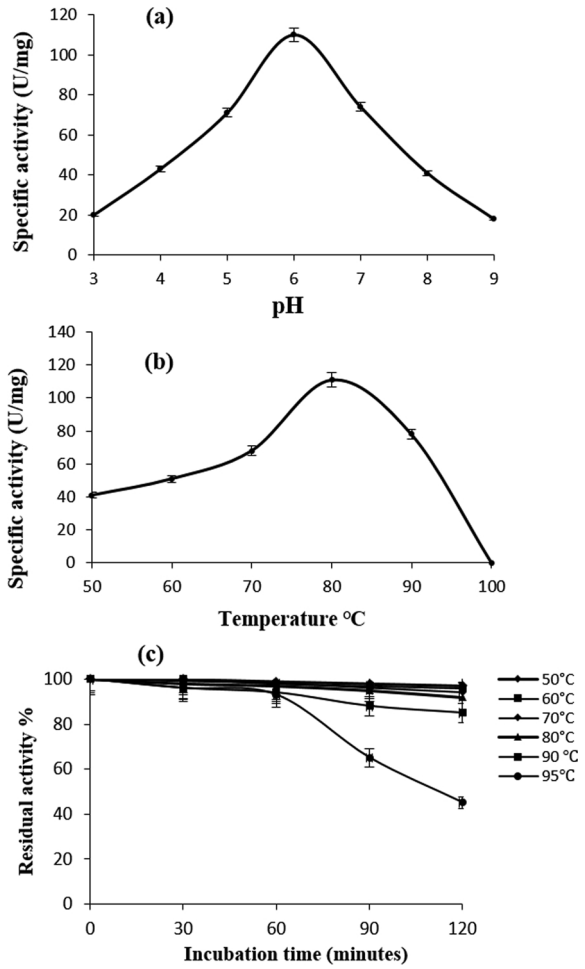

Fig. 3. Effect of temperature **(a)** and pH **(b)** on the activity of C6 mutant, and its stability at different temperatures **(c)**. The error bars represent the mean of three set of experiments with standard deviation < 5%.

GH11 from *Thermobacillus xylanolyticus*, which on directed evolution resulted in improved binding for birchwood xylan and thus several fold increased activity (Song et al., 2012).

3.4. Hydrolysate analysis

HPLC analysis of the sugars released from both the soluble substrate (CMC) and the insoluble substrates such as RAC and Avicel by the action of C6 mutant showed that cellobiose is the major product in the hydrolysate (Fig. 5a). This result suggests that C6 mutant of Cel12A is a cellobiohydrolase (Fig. 5b–d), which is contrary to a previous report claiming that Cel12A of *T. neapolitana* is an endoglucanase (Bok et al., 1998). Our results are supported by other studies which showed that cellobiose is the main product released from microcrystalline cellulose by the action of cellobiohydrolases (Liu et al., 2011). Cellobiohydrolase

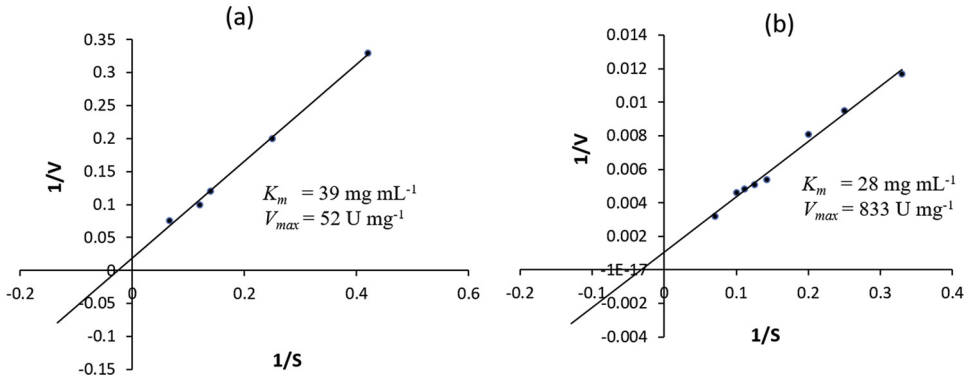

Fig. 4. Lineweaver-Burk plot for determination of the K_m and V_{max} values of the wild type **(a)** and C6 mutant **(b)**.

Table 2
Binding affinity of the wild type Cel12A and itsimproved mutant towards pretreated rice straw.

Variants	Total activity U/mL	Unbound activity U/mL	Unbound activity (%)	Binding efficiency (%)
Wild type	7.4 ± 0.95	4.8 ± 0.14	59.4	40.6
C6	23.9 ± 0.5	12.0 ± 0.25	50.2	49.8

(CBH) cleaves the β (1–4) glycosidic linkages by a double-displacement mechanism, resulting in release of cellobiose (Divne et al., 1994).

3.5. Protein-ligand docking

The 3-dimensional (3D) structures of the wild type Cel12A of *T. neapolitana* and its C6 mutant are predicted by the I-TASSER pipeline with a confidence score (C-score) of 1.33 and 1.21. Here, C-score has a value in [−5, 2] and benchmark data shows that a model with C-score above -1.5 should have a correct fold (Zhang, 2008). The high C-scores for Cel12A and the mutant suggest the highly likelihood of the correctness of the I-TASSER models.

Template based modeling approach (TBM) includes both homology modeling and threading can be used for the highly homologous protein targets. We used I-TASSER because it is a hierarchical method based on multiple threading template alignments. One of the major advantages of I-TASSER over traditional TBM approaches is that the fragment assembly simulations can often draw the template structures closer to the native.

I-TASSER models have shown higher modeling accuracy in CASP experiments than the template-based approaches (*i.e.*, in the category of easy TBM targets). This superiority of performance was mainly due to I-TASSER's ability for template structure refinement for the TBM targets. Fig. 6b and d show that in both the wild type and C6 mutant, the residues R60, Y65, E116, W118, E134, M136 and E231 are involved in the substrate interaction, where all these residues are located in a catalytic cavity consisting of β-sheets. E134 and E231 have been reported as active site residues of Cel12A (PDB ID: E134C) of *T. maritima* (Cheng et al., 2011) which has > 86% similarity in amino acid sequence with Cel12A of *T. neapolitana.* Previous study also confirmed the two glutamate residues to be essential for the catalytic activity of GH12 family cellulases. R60 was reported to be involved in making hydrogen bonds with the sugar residues in cellulose substrate at C1 (Cheng et al., 2011).

Wild type Cel12A model was docked with cellotriose ligand using EDock server. Binding energy between the ligand and protein was calculated in term of X-Score, which is the energy function of EDock. X score is the cumulative score of electrostatic energies (van der Waal, hydrogen bonding and hydrophobic forces). The X-Score showed by wild type Cel12A and C6 mutant was 4.8 and 5.6, respectively.

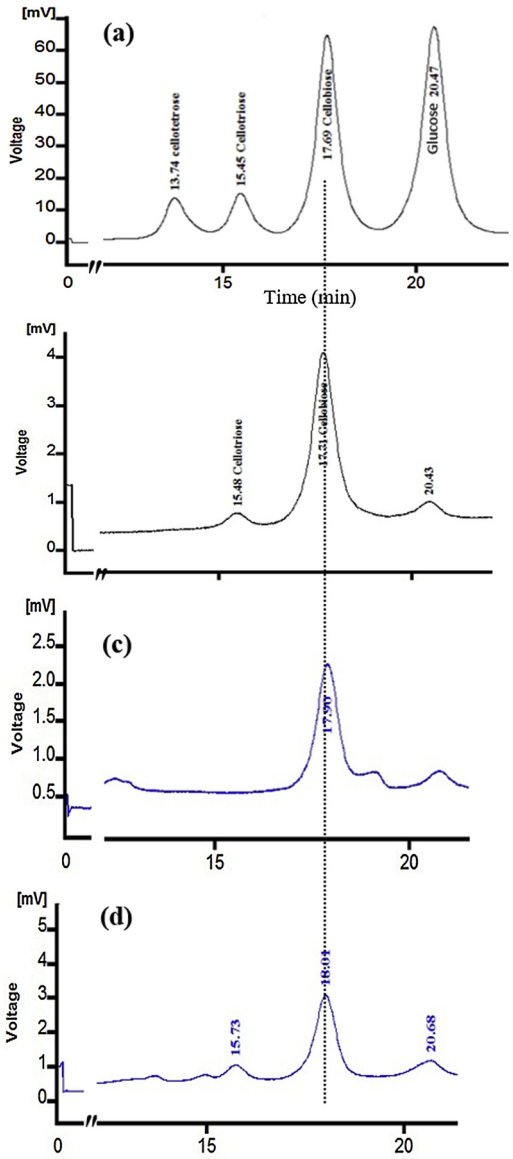

Fig. 5. HPLC analysis of the sugars obtained on hydrolysis of different substrates by C6 mutant of Cel12A. **(a)** Standard mixture of the sugars; **(b)** CMC; **(c)** RAC; **(d)** Avicel.

Computational model of Cel12A and C6 mutant shows the binding geometry between the catalytic residues and cellulose ligand (Fig. 6a and c).

The protein-ligand interactions of both the wild type and C6 mutant

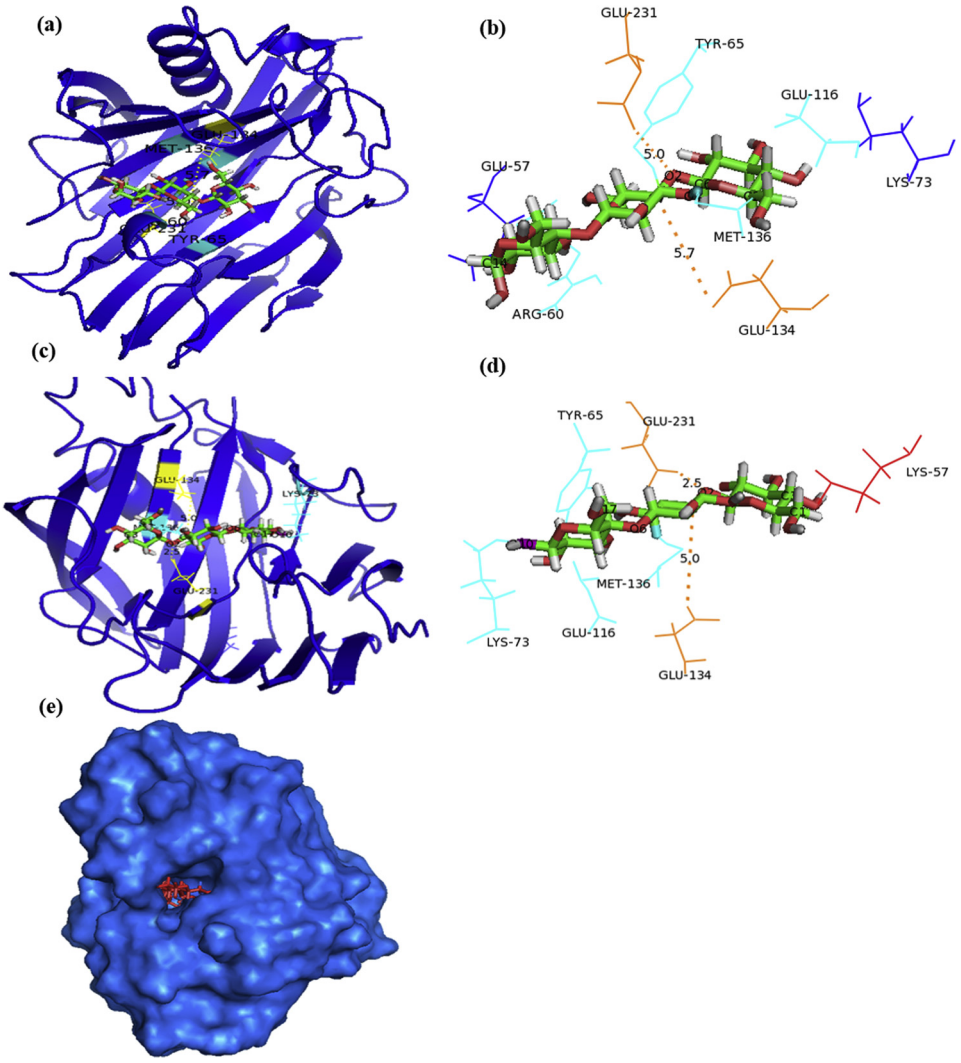

Fig. 6. Predicted structures of the wild type Cel12A **(a)** and its C6 mutant **(c)** docked with cellotriose ligand. In wild type the binding residues R65 and M136 (cyan) are involved in ligand binding **(b)**. However, in case of the mutant, K73 in addition to R65 and M136 is also involved in ligand binding. The distance between the active site residues E134 and E231 (orange) and the ligand for the wild type and the mutant, measured in A°, are indicated with orange dash lines. The distance between the ligand and E231 of the mutant are decreased to half than that of the wild type. The mutation K57 (red) located in the proximity of catalytic interface and due to its similarity of the side chain with that of R60 and K73 may enhance the ligand binding affinity of the mutant enzyme**(d)**. Surface model showing the tunnel shaped structure in complex with ligand **(e)** (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

studied through Edock server showed that the nucleophile Glu134 interact with anomeric carbon C7 and distort the glycosidic linkage; this allows the Glu231 to interact with O2 located in its close proximity and transfer the proton to glycosidic oxygen, thus releasing cellobiose. The reaction mechanism is similar to cellulose hydrolysis mechanism reported previously (Knott et al., 2013). The binding residues M136 and Y65 bind to H18 and H11 atoms of ligand, respectively, thus promoting the transition state. However, in C6 mutant an additional binding residue K73 bind to O10 of the ligand thus, enhancing the binding affinity of enzyme towards the substrate (Fig. 6d).

We noted that the distance between the active site residues of Cel12A, Glu134 and Glu231 and the ligand atoms 7C and O2 was 5.7 and 5.0 Å, respectively (Fig. 6b). However, in the case of C6 mutant, the distance between Glu231 and O2 reduced to 2.5 and 5 Å for Glu134 and 7C, respectively (Fig. 6d). The reduction in distance of the active site residues and ligand atoms may enhance the catalytic efficiency of the mutant.

The mutation E57 K obtained in the last round of ep-PCR is located in the catalytic cavity in close proximity to R60. The functional group of lysine, like arginine, protrudes towards the substrate conspicuously in the catalytic cavity. The mutants with V88A, S157 P, H165 K have the altered residues located away from catalytic cavity on the surface of the protein structure may have no direct interaction with the substrate. The proline and the alanine residues positioned at the surface may enhance hydrophobicity of the protein, which may contribute in enhancing the activity due to formation of more compact enzyme substrate complex under the aqueous conditions. These results are similar to those previously reported, where glycine and valine positioned at surface of the protein structure enhanced the enzyme activity due to increasing the hydrophobicity of protein (Zhang et al., 2015).

The surface model of both the wild type and the mutant showed presence of a catalytic tunnel (Fig. 6e) typical of cellobiohydrolases, which preferably release cellobiose (Divne et al., 1994). The cellulose chain is threaded through its reducing end into the catalytic tunnel with entry of two glucosyl residues at a time, followed by depolymerization of the chain by releasing free cellobiose units (Divne et al., 1994; Igarashi et al., 2009). The positioning of E134 and E231 above and below, respectively, in the catalytic tunnel of both the wild type Cel12A and C6 mutant attacks the acetal linkage by forming hydrogen bonds to hydroxyl groups of the glucosyl unit in cellulose chain and preferably releases cellobiose. These structural features strengthen our claim that the Cel12A have characteristics similar to those of the cellobiohydrolases.

4. Conclusion

Cel12A from *T. neapolitana* was optimized for their properties through directed evolution using ep-PCR. The best evolved mutant C6 with five mutations (N47S, E57 K, V88A, S157 P and K165 H) showed 5.0−, 4.8− and 2.7−fold increase in specific activity against RAC, Avicel and CMC, respectively, compared to that of the wild type enzyme. Protein modeling and docking analysis of C6 mutant showed that

the mutation E57K plays a key role in enhancing the enzyme activity by producing conformational changes in the catalytic cavity. HPLC analysis of the hydrolysate showed the release of cellobiose from cellulose as the major product and the possession of catalytic tunnel in the structure suggests that Cel12A is a cellobiohydrolase.

Declaration of Competing Interest

The authors have no conflicts of interest associated with this research work.

Acknowledgments

This research work was funded by the Higher Education Commission, Pakistan (Grant number 0.20-2765/NRPU/R&D/HEC/13). We are also thankful to IRSIP, HEC for facilitating the computational study carried out in Prof. Yang Zhang Lab, DCMB and University of Michigan, USA.

References

Ahmad, S., Ma, H., Akhtar, M.W., Zhang, Y.-H.P., Zhang, X.-Z., 2014. Directed evolution of *Clostridium phytofermentans* glycoside hydrolase family 9 endoglucanase for enhanced specific activity on solid cellulosic substrate. Bioenergy Res. 7, 381–388.

Basit, A., Akhtar, M.W., 2018. Truncation of the processive Cel5A of *Thermotoga maritima* results in soluble expression and several fold increase in activity. Biotechnol. Bioeng. 115, 1675–1684.

Bok, J.-D., Yernool, D.A., Eveleigh, D.E., 1998. Purification, characterization, and molecular analysis of thermostable cellulases CelA and CelB from Thermotoga neapolitana. Appl. Environ. Microbiol. 64, 4774–4781.

Cadwell, R.C., Joyce, G.F., 1992. Randomization of genes by PCR mutagenesis. Genome Res. 2, 28–33.

Cheng, Y.-S., Ko, T.-P., Wu, T.-H., Ma, Y., Huang, C.-H., Lai, H.-L., Wang, A.H.-J., Liu, J.-R., Guo, R.-T., 2011. Crystal structure and substrate-binding mode of cellulase 12A from Thermotoga maritima. Proteins: Struct., Funct., and Bioinf. 79, 1193–1204.

Chen, Z., Pereira, J.H., Liu, H., Tran, H.M., Hsu, N.S.Y., et al., 2013. Improved activity of a thermophilic cellulase, Cel5A, from *Thermotoga maritima* on ionic liquid pretreated switchgrass. PLoS One 8 (11), e79725.

Croft, A.K., Chen, Z., Pereira, J.H., Liu, H., Tran, H.M., Hsu, N.S.Y., Dibble, D., Singh, S., Adams, P.D., Sapra, R., Hadi, M.Z., Simmons, B.A., Sale, K.L., 2013. Improved activity of a thermophilic cellulase, Cel5A, from *Thermotoga maritima* on ionic liquid pretreated switchgrass. PLoS One 8, e79725.

Dave, B.R., Sudhir, A.P., Subramanian, R.B., 2015. Purification and properties of an endoglucanase from Thermoascus aurantiacus. Biotechnol. Rep. Amst. (Amst) 6, 85–90.

Delano, W.L., 2002. The PyMOL molecular graphics system. Proteins: Struct., Funct., Bioinf. 30, 442–445.

Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J., Teeri, T., Jones, T., 1994. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from Trichoderma reesei. Science 265, 524–528.

Farrow, M.F., Arnold, F.H., 2011. High throughput screening of fungal endoglucanase activity in Escherichia coli. J. Vis. Exp e2942-e2942.

Hong, J., Tamaki, H., Yamamoto, K., Kumagai, H., 2003. Cloning of a gene encoding thermostable cellobiohydrolase from *Thermoascus aurantiacus* and its expression in yeast. Appl. Microbiol. Biotechnol. 63, 42–50.

Igarashi, K., Koivula, A., Wada, M., Kimura, S., Penttilä, M., Samejima, M., 2009. High speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. J. Biol. Chem. 284, 36186–36190.

Kaur, J., Sharma, R., 2006. Directed evolution: an approach to engineer enzymes. Crit. Rev. Biotechnol. 26, 165–199.

Khan, M.I.M., Sajjad, M., Sadaf, S., Zafar, R., Niazi, U.H.K., Akhtar, M.W., 2013. The nature of the carbohydrate binding module determines the catalytic efficiency of xylanase Z of Clostridium thermocellum. J. Biotechnol. 168, 403–408.

Knott, B.C., Haddad Momeni, M., Crowley, M.F., Mackenzie, L.F., Götz, A.W., Sandgren, M., Withers, S.G., Ståhlberg, J., Beckham, G.T., 2013. The mechanism of cellulose hydrolysis by a two-step, retaining cellobiohydrolase elucidated by structural and transition path sampling studies. J. Am. Chem. Soc. 136, 321–329.

Liang, C., Fioroni, M., Rodríguez-Ropero, F., Xue, Y., Schwaneberg, U., Ma, Y., 2011. Directed evolution of a thermophilic endoglucanase (Cel5A) into highly active Cel5A variants with an expanded temperature profile. J. Biotechnol. 154, 46–53.

Ling, L., Chenggen, F., Weiqian, H., 2016. Improving the activity of the endoglucanase, Cel8M from *Escherichia coli* by error-prone PCR. Enzyme. Microb. Tech. 86, 52–58.

Liu, Y.-S., Baker, J.O., Zeng, Y., Himmel, M.E., Haas, T., Ding, S.-Y., 2011. Cellobiohydrolase hydrolyzes crystalline cellulose on hydrophobic faces. J. Biol. Chem., jbc. M110., 216556.

Nakazawa, H., Okada, K., Onodera, T., Ogasawara, W., Okada, H., Morikawa, Y., 2009. Directed evolution of endoglucanase III (Cel12A) from Trichoderma reesei. Appl. Microbiol. Biotechnol. 83, 649–657.

Pritchard, L., Corne, D., Kell, D., Rowland, J., Winson, M., 2005. A general model of error-prone PCR. J. Theor. Biol. 234, 497–509.

Rehbein, P., Harald, S., 2015. Integrated protocol for reliable and fast quantification and documentation of electrophoresis gels. Protein. Expression. Purif. 110, 1–6.

Rubin-Pitel, S., Cho, C., Chen, W., Zhao, H., 2006. Directed evolution tools in bioproduct and bioprocess development. Bioprocessing for Value-Added Products from Renewable Resources: New Technologies and Applications. pp. 49–72.

Sajjad, M., Khan, M.I.M., Akbar, N.S., Ahmad, S., Ali, I., Akhtar, M.W., 2010. Enhanced expression and activity yields of *Clostridium thermocellum* xylanases without non-catalytic domains. J. Biotechnol. 145, 38–42.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675.

Shi, H., Wang, L., Li, X., Wang, L., Zhang, Y., Li, X., Wang, F., 2014a. Directed Evolution of a hyperthermophilic endoglucanase Cel12B from *Thermotoga maritima*. BioResources 9, 3526–3535.

Song, L., Siguier, B., Dumon, C., Bozonnet, S., O'Donohue, M.J., 2012. Engineering better biomass-degrading ability into a GH11 xylanase using a directed evolution strategy. Biotechnol. Biofuels 5, 3.

Wang, T., Liu, X., Yu, Q., Zhang, X., Qu, Y., Gao, P., Wang, T., 2005. Directed evolution for engineering pH profile of endoglucanase III from Trichoderma reesei. Biomol. Eng. 22, 89–94.

Wang, Y., Wang, X., Tang, R., Yu, S., Zheng, B., Feng, Y., 2010. A novel thermostable cellulase from Fervidobacterium nodosum. J. Mol. Catal., B Enzym. 66, 294–301.

Xiao, Z., Storms, R., Tsang, A., 2005. Microplate-based carboxymethylcellulose assay for endoglucanase activity. Anal. Biochem. 342, 176–178.

Yang, J., Roy, A., Zhang, Y., 2013. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 29, 2588–2595.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2015. The I-TASSER Suite: protein structure and function prediction. Nat. Methods 12, 7–8.

Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9, 40.

Zhang, Y.H.P., Zhang, J., Shi, H., Xu, L., Zhu, X., Li, X., 2015. Site-Directed mutagenesis of a hyperthermophilic endoglucanase Cel12B from *Thermotoga maritima* based on rational design. PLoS One 10, e0133824.

Zhang, X.Z., Zhang, Z., Zhu, Z., Sathitsuksanoh, N., Yang, Y., Zhang, Y.H., 2010. The noncellulosomal family 48 cellobiohydrolase from *Clostridium phytofermentans* ISDg: heterologous expression, characterization, and processivity. Appl. Microbiol. Biotechnol. 86, 525–533.