J. Microbiol. Biotechnol. (2007), **17**(8), 1249–1253

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

Analysis of Active Center in Hyperthermophilic Cellulase from *Pyrococcus horikoshii*

KANG, HEE-JIN AND KAZUHIKO ISHIKAWA*

National Institute of Advanced Industrial Science and Technology (AIST, Kansai), 1-18-31, Midorigaoka, Ikeda, Osaka 563-8577, Japan

Received: January 6, 2007 Accepted: March 23, 2007

Abstract A hyperthermostable endoglucanase from *Pyrococcus horikoshii* with the capability of hydrolyzing crystalline cellulose was analyzed. A protein engineering study was carried out to obtain a reduced-size mutant. Five amino acid residues at both the N- and C-terminus were found to be removable without any loss of activity or thermal stability. Site-directed mutagenesis was also performed on R102, N200, E201, H297, Y299, E342, and W377, residues possibly involved in the active center or in the recognition and binding of a cellulose substrate. The activity of the resulting mutants was considerably decreased, confirming that the mutated residues were all important for activity. A reduced-size enzyme, as active as the wild-type endoglucanase, was successfully obtained, plus the residues critical for its activity and specificity were confirmed. Consequently, an engineered enzyme with a reduced size was obtained, and the amino acids essential for activity were confirmed by site-directed mutagenesis and comparison with a known three-dimensional structure.

Keywords: Endoglucanase, hyperthermophile, compact enzyme, active site, site-directed mutagenesis

Cellulose is the most abundant biomass and hence the most potentially important renewable carbon resource. An essential, yet not established step in biochemically recycling cellulose is the hydrolysis of solid cellulose into soluble sugars by cellulolytic enzymes that cleave the β-1, 4 linkages between each glucose unit. The complete hydrolysis of cellulose requires several types of enzyme: endo-1,4-β-glucanase (EC3.2.1.4), 1,4-β-D-cellobiohydrolase (EC3.2.1.9), and 1,4-β-glucosidase (EC3.2.1.21) [12]. Various endoglucanases from cellulolytic microorganisms, including fungi and bacteria, have already been investigated.

*Corresponding author
Phone: 81-72-751-9526; Fax: 81-72-751-9628;
E-mail: kazu-ishikawa@aist.go.jp

The present authors recently reported on a hyperthermostable endoglucanase (family 5) from *Pyrococcus horikoshii* (EGPh) that can hydrolyze crystalline cellulose, which resists the action of most other cellulases [1, 3, 4]. As such, this enzyme is highly promising for industrial application, where the emergence of a hyperthermostable endoglucanase is awaited to complement the existing hyperthermostable amylase. Furthermore, a system for mass production of the enzyme on a kg-t scale has already been successfully established by a JST-assisted project (http://www.jst.go.jp/pr/info/info155/index.html) in Japan. Accordingly, with industrial application in mind, the present authors initiated various studies on the structure–function relationship of EGPh with the hope of potentially improving the enzyme's action, such as its specificity, thermal stability, and tolerance to extreme pH. Yet, despite several attempts to crystallize EGPh and determine its three-dimensional structure, the crystallization was found to be a difficult task. Thus, biochemical methods were used to explore the action mechanism of EGPh and identify possible ways to improve its action.

The amino acid sequence of EGPh exhibits a 43% homology with that of the endoglucanase from *Acidothermus cellulolyticus*, whose three-dimensional structure has already been solved [1, 9]. Using this crystal structure data, Sakon *et al.* [9] have previously speculated on the importance of several amino acids in binding and recognizing the cellulose substrate, in addition to the two carboxyl residues that are directly involved in hydrolysis. Therefore, since the high similarity between the two endoglucanases in the amino acid sequence could suggest a similar importance for the corresponding amino acid residues in EGPh, the present study constructed mutant enzymes employing the candidate amino acid residues. Furthermore, a reduced-size EGPh was constructed based on removing the unstable N-terminal and C-terminal regions.

MATERIALS AND METHODS

Construction of Reduced-Size Cellulase Starting from EGPh

To construct a reduced-size EGPh, the enzyme was first digested with a protease, bovine pancreatic α-chymotrypsin (Wako Pure Chemical Industries, Osaka, Japan), where the EGPh was incubated with α-chymotrypsin at various enzyme/substrate ratios for 12 h at 37°C [10]. The results of the proteolysis were analyzed by SDS-PAGE, and the electrophoretic band cut out for N-terminal amino acid sequencing, which was performed using an HP G1005A protein sequencing system (APRO, Naruto, Tokushima, Japan). Three kinds of mutant enzyme were then constructed based on deleting five amino acid residues at the N-terminus, five amino acid residues at the C-terminus, or ten amino acid residues at the C-terminus, designated as ΔN5, ΔC5, and ΔC10, respectively. Furthermore, ΔN5C5 was constructed based on deleting five residues at both termini.

Mutation of Active-Site Candidates

Site-directed mutagenesis was performed on various active-site candidate positions through analyzing the structure of EGPh and *A. cellulolyticus*. Sakon *et al.* [9] previously suggested that the eight key residues of R62, N161, E162, H238, Y240, E282, W319, and D327 were the conserved residues in *A. cellulolyticus*, which in the present study were determined to correspond to R102, N200, E201, H297, Y299, E342, W377, and D385, respectively, in EGPh. Therefore, the mutagenesis was based on this information and carried out using a PCR. The PCR was applied using wild-type EGPh as the template [7], and resulted in the construction of the mutants E201Q, H297N, Y299F, E342Q, and D385N. Furthermore, the residues R102, N200, E201, H297, Y299, E342, and W377 were also changed to Ala.

Expression of Mutant Proteins

The resulting mutated DNA fragment was digested with NdeI and BamHI and inserted into the corresponding site of vector pET11a (Novagen, Madison, WI, U.S.A.). All the vectors carrying the mutations were then transformed into *E. coli* XL-2 Blue (Stratagene, La Jolla, CA, U.S.A.), and the constructed plasmids introduced into *E. coli* strain BL21(DE3)pLysS. Thereafter, each transformant was cultured in an LB broth containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol at 37°C until the OD600 reached 0.2, and then isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added (0.01 mM). Purification of the recombinant enzymes was carried out using a method reported previously [4].

Enzyme Activity and Optimal pH Measurements

The hydrolytic activity of the enzymes toward carboxy methyl cellulose (CMC) (Wako) was measured using the modified Somogi-Nelson method [2] at 85°C in a 100 mM acetate buffer (pH 5.6). The activity against *p*-nitrophenyl cellobiose (G2-PNP) (Sigma, St. Louis, MO, U.S.A.) was measured in a 100 mM acetate buffer (pH 5.6) from the absorbance at 420 nm of *p*-nitrophenol (PNP) liberated from G2-PNP at 50°C. G2-PNP is a useful water-soluble substrate analog for characterizing cellulase [8]. Additionally, K_m and k_{cat} were determined using the Henri-Michaelis-Menten equation [5] and nonlinear least squares method [11]. The optimal pH for the enzymes was examined using CMC as the substrate, and determined at 85°C using different pH values ranging from 4 to 10, where the buffers used for this purpose were as follows: pH 4–6 (50 mM acetate buffer), pH 7–8 (50 mM phosphate buffer), and pH 9–10 (50 mM borate buffer).

Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) measurements of the ΔN5, ΔC5, and ΔN5C5 mutants were carried out with a nanoDSCII instrument (Calorimetry Sciences, Spanish Fork, UT, U.S.A.) using platinum tubing cells with a volume of 0.3 ml. The proteins were dialyzed against a 50 mM phosphate buffer, pH 7.0. The dialysate buffer was also used as the reference solution for the DSC scan. Samples containing 1.0 mg/ml of protein were heated at 1°C/min from 0 to 125°C.

Thermostability Evaluated Through Activity

The thermostability of ΔN5, ΔC5, and ΔN5C5 was determined by incubating the enzymes for various times (0–6 days) at 90°C. After cooling the enzymes on ice for 10 min, the residual activity toward 0.5% CMC was determined using the modified Somogi-Nelson method [2] at 85°C in a 100 mM acetate buffer (pH 5.6).

RESULTS

Preparation of Reduced-Size Enzymes

To remove the flexible part of the molecule, EGPh was proteolyzed using α-chymotrypsin, and the size of EGPh (43 kDa) decreased to approximately 40 kDa, as shown in Fig. 1. The digested enzyme was stable at room temperature. To identify the deleted amino acid residues at the N-terminus, the digested enzyme (40 kDa) was analyzed using a protein sequencing system, and the first five amino acid residues at the N-terminus of the digested enzyme were found to be QTPTG, indicating the removal of the five residues ENTTY from the N-terminus of EGPh. Thus, a mutant EGPh gene lacking the first five amino acids at the N-terminus (ΔN5) was constructed. The analysis of the C-terminus amino acid residues in the digested enzyme was more difficult, and therefore, five and ten residues at the C-terminus (ΔC5 and ΔC10) of EGPh were removed.

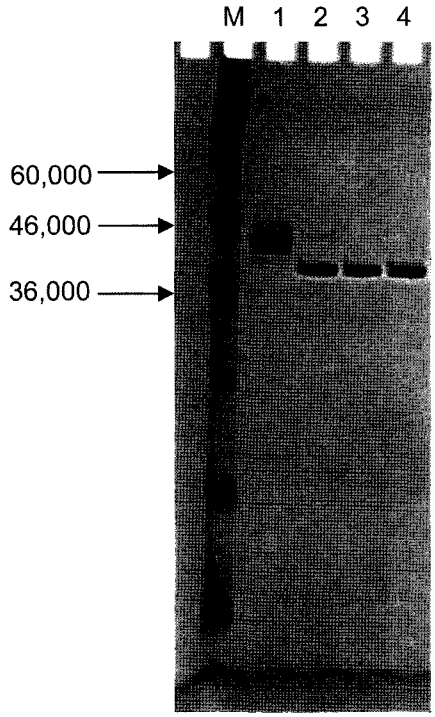

Fig. 1. Proteolysis of wild-type EGPh by α-chymotrypsin at 37°C for 12 h. The digested mixture was analyzed by SDS-PAGE.
Lane 1, without proteolysis; lane 2, S/E (substrate vs enzyme) ratio=3,200; lane 3, S/E ratio=1,600; lane 4, S/E ratio=800.

Thus, mutant EGPh genes lacking five and ten residues at the C-terminus were constructed. Additionally, their recombinant enzymes (ΔN5, ΔC5, and ΔC10) were prepared and purified according to the Materials and Methods above.

Table 1. Relative activities of recombinant enzymes toward CMC.

Enzymes	Relative activity* (%)
WT	100
ΔN5	135.6
ΔC5	111.6
ΔN5C5	111.1
E201Q	1.12
H297N	1.31
Y299F	2.15
E342Q	0.01
D385N	29.9
R102A	0.67
N200A	5.43
E201A	0.01
H297A	0.08
Y299A	0.21
E342A	0.08
W377A	1.02

*Relative activity against the wild-type (WT) EGPh activity as the standard.

Table 2. Activities of recombinant enzymes toward G2-PNP.

Enzymes	K_m (mM)	k_{cat} (1/s)	k_{cat}/K_m
WT	0.95±0.33	0.41±0.05	0.43±0.16
ΔN5	0.76±0.17	0.78±0.04	1.02±0.24
ΔC5	1.44±0.33	0.75±0.04	0.52±0.12
ΔN5C5	1.13±0.19	0.91±0.00	0.79±0.13
E201Q	1.24±0.75	0.01±0.00	0.01±0.01
H297N	2.76±0.62	0.01±0.00	0.00±0.00
Y299F	1.12±0.42	0.01±0.00	0.01±0.00
E342Q	N.D.	N.D.	N.D.
D385N	2.40±0.78	0.11±0.01	0.04±0.01

The concentration of *p*-nitrophenol was calculated in comparison with the standard curve of a *p*-nitrophenol solution in a 100 mM acetate buffer (pH 5.6). The k_{cat} values were determined from the initial velocities measured in the presence of excess concentrations of the substrates.

The hydrolytic activity toward CMC was slightly increased for ΔN5 and ΔC5, yet decreased to 0.32% for ΔC10 (Table 1). Consequently, it was decided to delete five residues from each end, resulting in ΔN5C5, to engineer a reduced-size enzyme that was expected to possess thermostability without any loss of activity. The activity of ΔN5C5 was determined as 111% compared with the activity of the wild-type (WT) (Table 1).

The kinetic parameters of the enzymes determined using G2-PNP are shown in Table 2. The kinetic parameters for G2-PNP with ΔN5, ΔC5, and ΔN5C5 were slightly increased when compared with those for the WT.

DSC measurements were performed for the WT, ΔN5, ΔC5, and ΔN5C5 enzymes, and the peak temperatures and ΔH for heat denaturation obtained from thermograms

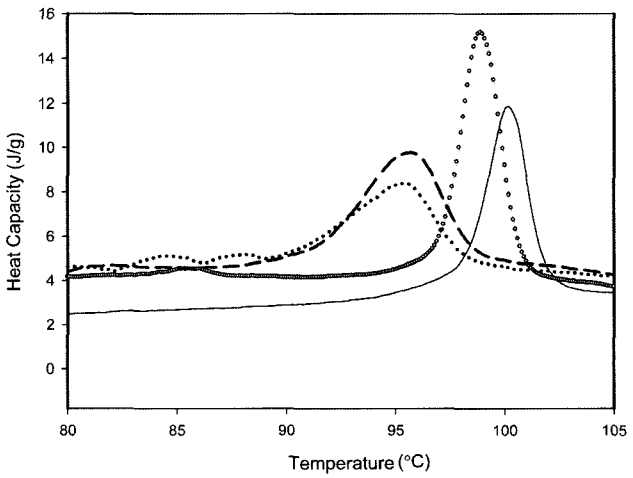

Fig. 2. Differential scanning calorimetry thermograms of wild-type and recombinant enzymes.
The protein concentration was 1 mg/ml in 50 mM sodium phosphate, pH 7. The T_m values of WT, ΔN5, ΔC5, and ΔN5C5 were 100.1, 98.9, 95.5, and 95.3°C, respectively. The ΔH values of WT, ΔN5, ΔC5, and ΔN5C5 were 23, 25, 21, and 17 J/g, respectively. (——) WT, (–○–) ΔN5, (·····) ΔC5, (— —) ΔN5C5

(Fig. 2). It was found that the peak temperature for the wild-type (100°C) was not that different (95–99°C) for the three mutants. In addition, the ΔH of 23 J/g for the wild-type was also similar to those for the mutants (17–25 J/g). Thus, the thermostability of the enzyme was not significantly influenced by the removal of the residues.

The irreversible thermostability was also evaluated by incubation for various time periods (0–6 days) at 90°C, followed by determination of the residual activity toward 0.5% CMC. The time needed for total activity loss was 2, 3, 4, and 6 days for the ΔN5, ΔC5, WT, and ΔN5C5 enzymes, respectively. Thus, the irreversible thermostability of ΔN5C5 was even longer than that of the WT.

Point Mutation of Candidate Active-Site Residues

Kashima *et al.* [4] considered, based on sequence comparison, that the essential residues for the hydrolysis of cellulose include E201 and E342. However, according to the discussion by Sakon *et al.* [9], certain regions adjacent to the active center are also potentially involved in the recognition and binding of the cellulose substrate. Thus, site-directed mutagenesis was performed on such residues to verify any changes in specificity/pH tolerance, and the results are presented in Tables 1 and 2. The activities toward CMC by E201Q, H297N, Y299F, and E342Q were in fact reduced to 1–2% of that by the WT, indicating that these positions are all important for activity. However, the activity of D385N was higher than expected, about 30% of that of the WT.

The kinetic parameters determined using G2-PNP are shown in Table 2. All the k_{cat} values for the mutant enzymes, except for D385N, decreased greatly, even by two orders of magnitude. However, for D385N, the reduction was less marked, only to about 1/4.

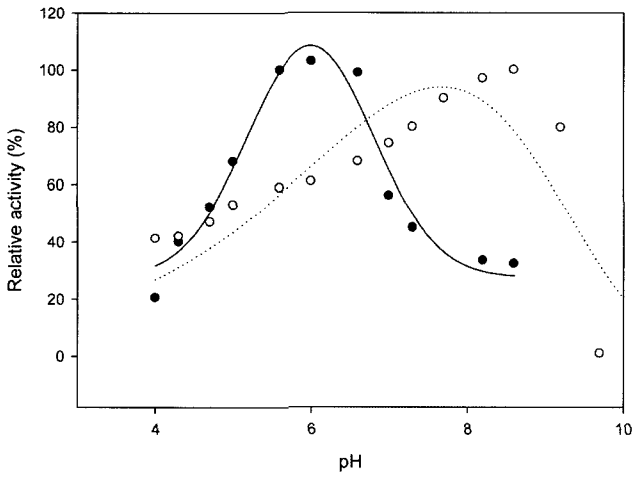

Fig. 3. pH activity-profile of WT (●—●) and R299F (○···○) represented as relative activities toward CMC.
The relative activities of WT and Y299F were calculated as a percentage against the activity at pH 5.6 and 8.6, respectively. The absolute activity value of the Y299F mutant at pH 5.6 corresponded to 2.15% of the wild-type EGPh activity.

The optimal pH values for the mutant enzymes were examined using CMC as the substrate. The optimal pH for E201Q, E342Q, and D385N was between 5.5 and 6.0, similar to that for the wild-type. However, the optimal pH values for H297N and Y299F were 7.0 and 8.5, respectively (Fig. 3).

In addition, when the candidate positions for the active center, R102, N200, E201, H297, Y299, E342, and W377 were changed to Ala, in all cases, the activity decreased considerably (to less than 1% of that for the wild-type, except for N200A, in which case the reduction was only to 5%) (Table 1).

DISCUSSION

A compact enzyme was successfully obtained from wild-type EGPh based on removing five amino acid residues from the N- and C-terminal ends without losing the thermostability and activity.

To demonstrate the decisive role of the conserved residues that would seem, based on a sequence comparison with the endoglucanase of *A. cellulolyticus*, to be important for the cellulose hydrolysis of EGPh, 12 mutations of eight conserved residues near the substrate-binding site in a family 5 endoglucanase were studied for their effects on the catalytic activity, substrate specificity, and optimal pH. The activity results toward CMC and G2-PNP revealed that the positions of E201, H297, Y299, and E342 were important for activity. However, the activity of D385N was higher than expected, indicating that D385 was not as important for the catalytic action as the other four candidate residues. An analysis of the optimal pH profile for the Y299F enzyme also provided further information on the role of Y299. The optimal pH values for the Y299F mutant produced the largest increase in the optimal pH, amounting to nearly 3 pH units, suggesting an important role for tyrosine protonation in the catalytic action of hydrolyzing CMC.

Furthermore, mutant enzymes were constructed by employing the candidate loop positions, R102, N200, H297, Y299, and W377, in addition to E201 and E342 in the active center. The activity of all these mutant enzymes decreased considerably. Interestingly, all seven positions correspond to the conserved residues of a family-5 endoglucanase, plus these residues are involved in interactions with glucose units or glycosidic bonds, implying their importance for substrate recognition and binding. The positions of the first and second conserved Glu of a family-5 endoglucanase, corresponding to E201 and E342 in EGPh, have already been identified as the proton donor and nucleophile, respectively, in a displacement reaction [6, 9]. In a previous study on the

structure of *A. cellulolyticus*, the three residues N200, E201, and E342 were thought to hydrogen bond directly to the ligand, R102 and Y299 to hydrogen bond to the nucleophile E342, and H297 to hydrogen bond to the acid/base E201 [9]. Additionally, W319 in *A. cellulolyticus* corresponding to W377 in EGPh also contacts the C3, C4, and C5 atoms of Glc3. Thus, when taking into account the important roles of the amino acid residues as discussed by Sakon *et al.* [9], it is natural that the mutation of these residues into Ala greatly reduced the activity toward CMC. This fact also underlines the correctness of the present inference that both endoglucanases are similar in structure and function.

Acknowledgment

This work was performed as part of the JSPS fellowship program supported by the Japan Society for the Promotion of Science.

REFERENCES

1. Ando, S., H. Ishida, Y. Kosugi, and K. Ishikawa. 2002. Hyperthermostable endoglucanase from *Pyrococcus horikoshii*. *Appl. Environ. Microbiol.* **68:** 430–433.
2. Hiromi, K., Y. Takahashi, and S. Ono. 1963. Kinetics of hydrolytic reaction catalyzed by crystalline bacterial α-amylase. The influence of temperature. *Bull. Chem. Soc. Jpn* **36:** 563–569.
3. Kang, H. J., K. Uegaki, H. Fukada, and K. Ishikawa. 2007. Improvement of the enzymatic activity of the hyperthermophilic cellulase from *Pyrococcus horikoshii*. *Extremophiles* **11:** 251–256.
4. Kashima, Y., K. Mori, H. Fukada, and K. Ishikawa. 2005. Analysis of the function of a hyperthermophilic endoglucanase from *Pyrococcus horikoshii* that hydrolyzes crystalline cellulose. *Extremophiles* **9:** 37–43.
5. McCarter, J. D. and S. G. Withers. 1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **4:** 885–892.
6. Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich. 1986. Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51:** 263–273.
7. Percival Zhang, Y. H., M. E. Himmel, and J. R. Mielenz. 2006. Outlook for cellulase improvement: Screening and selection strategies. *Biotechnol. Adv.* **24:** 452–481.
8. Sakon, J., W. S. Adney, M. E. Himmel, S. R. Thomas, and P. A. Karplus. 1996. Crystal structure of thermostable family 5 endocellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose. *Biochemistry* **35:** 10648–10660.
9. Uegaki, K., M. Shirakawa, T. Fujita, T. Taniguchi, and Y. Kyogoku. 1993. Characterization of the DNA binding domain of the mouse IRF-2 protein. *Protein Eng.* **6:** 195–200.
10. Wilkinson, G. N. 1961. Statistical estimations in enzyme kinetics. *Biochem. J.* **80:** 324–332.
11. Wood, T. M. 1985. Properties of cellulolytic enzyme systems. *Biochem. Soc. Trans.* **13:** 407–410.