Journal of Bioscience and Bioengineering
VOL. 127 No. 5, 554–562, 2019

www.elsevier.com/locate/jbiosc

Cloning, expression, and characterization of novel GH5 endoglucanases from *Thermobifida alba* AHK119

Takeshi Ohta,[1,*] Hitomi Horie,[1] Aina Matsu-ura,[1] and Fusako Kawai[2]

Department of Life Science, Faculty of Life and Environmental Sciences, Hiroshima Prefectural University, Shobara, Hiroshima 727-0023, Japan[1] *and Center for Fiber and Textile Science, Kyoto Institute of Technology, Kyoto 606-8181, Japan*[2]

Received 12 March 2018; accepted 16 October 2018
Available online 2 December 2018

Thermobifida alba **AHK119 exhibits sufficient filter paper-degradation activity in its culture supernatant. *AHK119-bMs* (1365 bp) and *AHK119-E5* (1425 bp), which encode novel GH5 family endoglucanases, were cloned from the genomic DNA of *T. alba* AHK119. AHK119-bMs and AHK119-E5 consisted of 454 and 474 amino acid residues, respectively, in which the catalytic domain (CD) and carbohydrate-binding module (CBM) were connected by an accessary module (linker region). The amino acid sequences of CD and CBM of AHK119-bMs were most identical to those of endo-β-mannanases (Man5As) from *Thermobifida fusca* TM51, *T. halotolerans* YIM90462, and *T. cellulosilytica* TB100. In contrast, the amino acid sequences of CD and CBM of AHK119-E5 were most identical to those of endo-1,4-β-glucanases (cellulases; Cel5As) from *T. fusca* and *T. halotolerans* YIM90462. However, the linker region of both the genes shared low identities with those of Man5As and Cel5As. AHK119-bMs showed broader specificities toward cellulosic substrates than Man5As, whereas AHK119-E5 showed higher activity toward insoluble cellulosic substrates than toward soluble ones, which was conflicting when compared with other Cel5As. In addition, AHK119-bMs and AHK119-E5 showed different requirements for metal ions from those of Man5As and Cel5As, respectively. Therefore, both the enzymes were identified as novel GH5 endoglucanases, and the accessary modules seemed to play important roles in their enzymatic properties.**

[**Key words:** *Thermobifida alba* AHK119; GH5 endoglucanase; Cellulase; Mannosidase; Bioethanol production; Material production; Bioprocessing; Biorefinery]

Cellulose is the most abundant biopolymer on earth (1) and plays an important role in the recycling of carbon fixed by photosynthesis. Furthermore, it is an important material in industry because it exists in various polymeric forms, such as amorphous, crystalline, and semi-crystalline, which can be useful for diverse purposes, such as in the production of paper, textile, and food additives. Cellulose is often utilized for bioethanol production, and its saccharification is highly efficient in second-generation bioethanol production. For cellulose extraction from lignocellulosic material containing various polymers, such as cellulose, hemicellulose, and lignin, it is necessary to combine multiple types of carbohydrate-active enzymes, including a variety of glycosyl hydrolases (GHs; EC3.2.1.xxx) (2). Endoglucanase randomly attacks β-glycosidic bonds in the amorphous regions of cellulose or hemicellulose fibers and opens sites for the subsequent action of cellobiohydrolase. Based on amino acid sequences and three-dimensional (3D) structures, endoglucanases are classified into 11 GH families (GH 5–12, 44, 45, 48, 51, and 74). Among the GH families, GH5 is the largest and most functionally diverse. Enzymes belonging to the GH5 family are generally composed of a catalytic domain (CD) formed by a triosephosphate isomerase (TIM)-barrel fold (3), carbohydrate-binding module (CBM) (4), and an accessory module (linker region) which connects CD and CBM (4). Furthermore, the functions of family GH5 enzymes are influenced by differences in their CBMs, leading to more than 20 experimentally verified activities on different polysaccharides. Although cellulose has a simple chemical structure (β-1,4-linked glucosyl residues), it is cross-linked with hemicellulose in plants and exhibits low solubility in aqueous solutions. Therefore, depolymerization of cellulose and hemicellulose requires at least three complementary enzyme activities (endo-1,4-β-glucanase, exo-1,4-β-glucanase, and β-glucosidase). Species of the genus *Thermobifida* are aerobic actinomycetes belonging to the family Nocardiopsaceae and are known to produce multiple extracellular enzymes, including several cellulases, responsible for the decomposition of cellulose and lignocellulose. The genus consists of four species: *Thermobifida fusca* (5–8), *T. cellulosilytica* (9), *T. halotolerans* (accession no. PRJNA293778) and *T. alba* (10). Except *T. alba*, all other species' genome sequences have previously been reported. In particular, *T. fusca* has been well studied as a model organism of cellulolytic bacteria that can utilize various plant cell wall polymers due to the presence of multiple cellulases, including three endoglucanases, two exoglucanases, and one endo/exocellulase (11–16). However, there is only one report documenting the presence of β-xylanase (17) in *T. alba*.

In this study, we found that *T. alba* AHK119 isolated from compost (18) possessed strong cellulase activity. We attempted to clone the endoglucanases responsible for cellulose hydrolysis. Two types of endoglucanases were PCR-amplified from the total genome of *T. alba* AHK119 based on the nucleotide sequence of CBM2 from

* Corresponding author. Tel./fax: +81 859 56 3434.
E-mail address: escomplexlab@gmail.com (T. Ohta).

1389-1723/$ – see front matter
https://doi.org/10.1016/j.jbiosc.2018.10.015

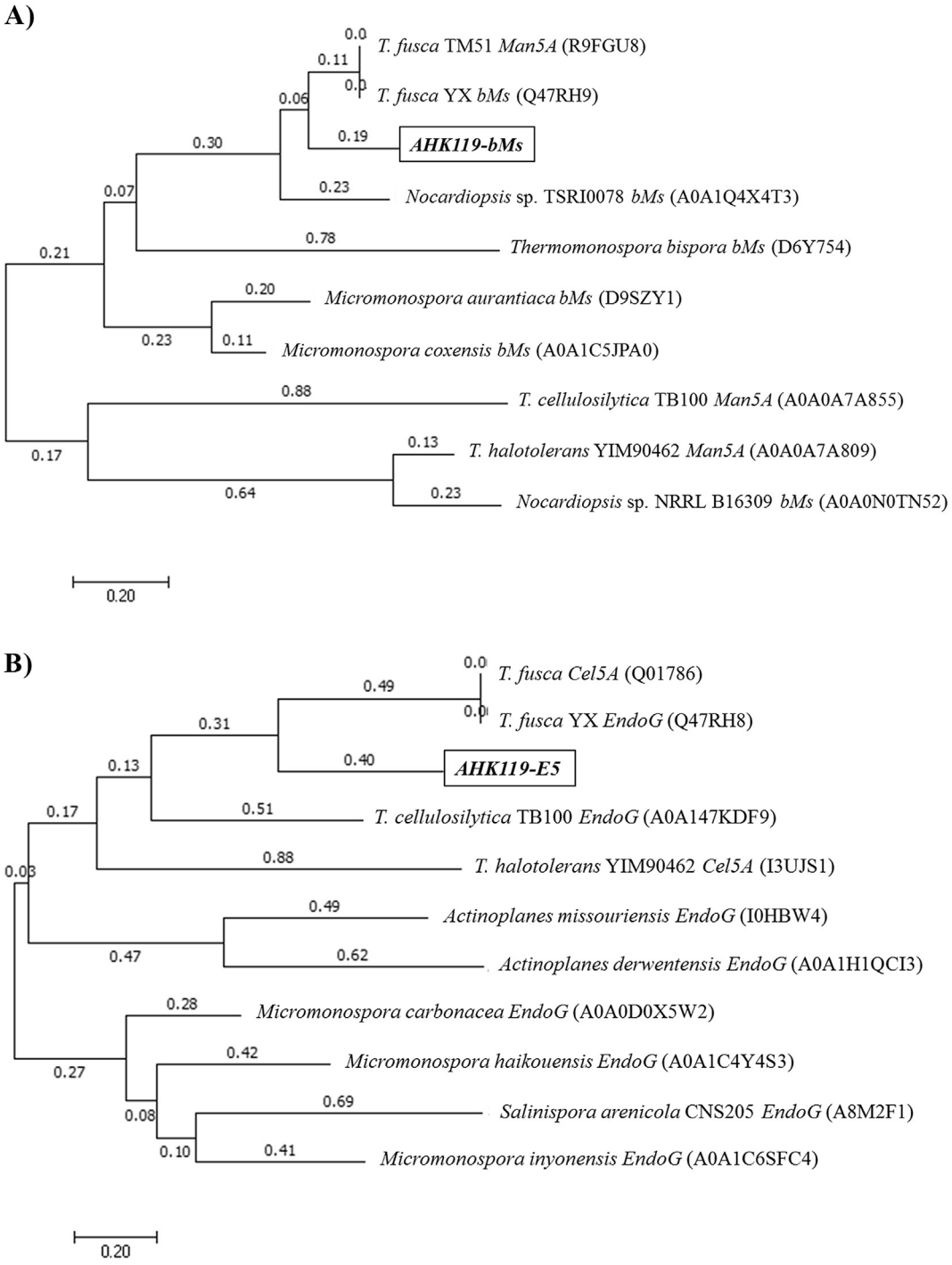

FIG. 1. (A) Molecular phylogenetic analysis of *AHK119-bMs* using the maximum likelihood method. (B) Molecular phylogenetic analysis of *AHK119-E5* using the maximum likelihood method.

T. fusca (10). We cloned and expressed two putative GH5 endoglucanases in *Escherichia coli*, which were characterized as novel GH5 endoglucanases. Their ability to degrade cellulose and hemicellulose and high thermal stability suggested their potential application in industrial processes.

MATERIALS AND METHODS

Microorganisms and culture conditions *T. alba* AHK119 strain was used as an original strain for cloning endoglucanases (18). It was cultured in Luria–Bertani (LB) medium at 50°C. *E. coli* DH5α (Toyobo, Osaka, Japan) and *E. coli* BL21(DE3) pLysS (Takara, Shiga, Japan) were used as cloning and expression hosts, respectively, and were grown in LB broth or agar supplemented with 50 μg/ml ampicillin at 37°C.

Gene cloning and expression of AHK119-bMs and AHK119-E5 Genomic DNA was isolated from *T. alba* AHK119 using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) following the manufacturer's instructions. *AHK119-bMS* was PCR-amplified using primers exbMs-N-BamHI (5′-GGATCCGACCATGAGAAAA-3′) and exbMs-C-HindIII (5′-AAGCTTTCAGTCGGTCGTG-3′) designed from β-mannosidase of *T. fusca* YX (Q47RH9), whereas *AHK119-E5* was PCR-amplified using primers exE5-N-BamHI (5′-GGATCCAACCATGACCAGACCC-3′) and exE5-C-HindIII (5′-AAGCTTTCAGCTGCGGAGCTG-3′) designed from endoglucanase of *T. fusca* YX (Q47RH8), with underlined sequences harboring restriction sites. PCR was performed using KAPA Taq EXtra DNA polymerase (Nippon Genetics, Tokyo, Japan) for 25 cycles of 15 s at 94°C, 20 s at 53°C, and 20 s at 70°C. The amplified

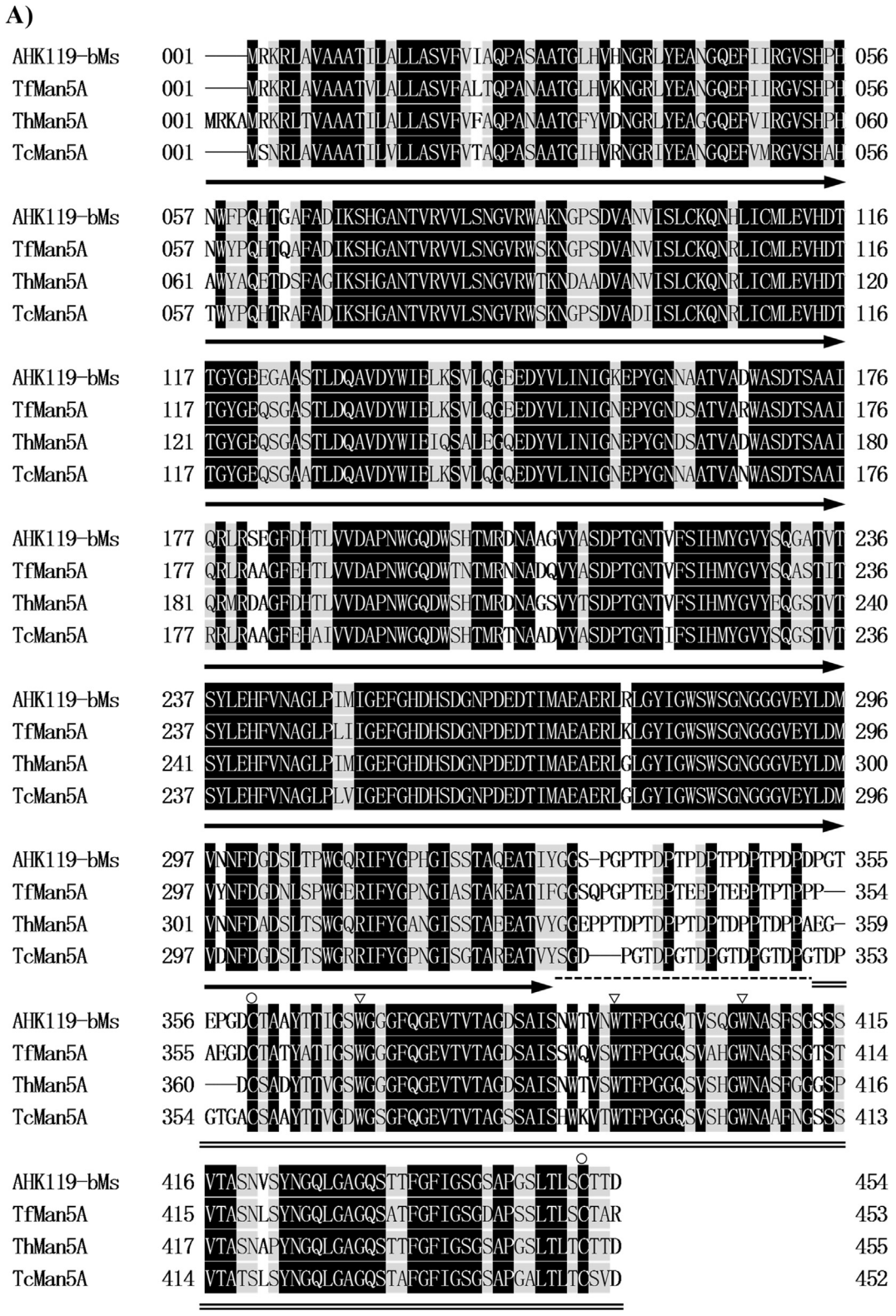

FIG. 2. (A) Sequence alignment of AHK119-bMs with Man5As. Open circle, cysteine residues related to a disulfide bond; open triangle, the surface-exposed tryptophan. TfMan5A, Man5A of *T. fusca* TM51 (R9FGU8); ThMan5A, Man5A of *T. halotolerans* YIM90462 (A0A0A7A809); and TcMan5A, Man5A of *T. cellulosilytica* TB100 (A0A0A7A855). (B) Sequence alignment of AHK119-E5 with Cel5As and endoglucanase. The CBM2a domain signature is indicated by a solid box. ThCel5A, Cel5A of *T. halotolerans* YIM90462 (I3UJS1); TcEndoG, endo-1,4 β-glucanase of *T. cellulosilytica* TB100 (A0A147KDF9); and TfCel5A, Cel5A of *T. fusca* (Q01786). Gaps, indicated by dashes, were introduced into the sequences to maximize homology. Identical amino acids are shown as white characters with black backgrounds and similar amino acids are shown with gray backgrounds. CD, the accessory module (linker region), and CBM2 are indicated by solid arrows, dotted lines, and double lines, respectively.

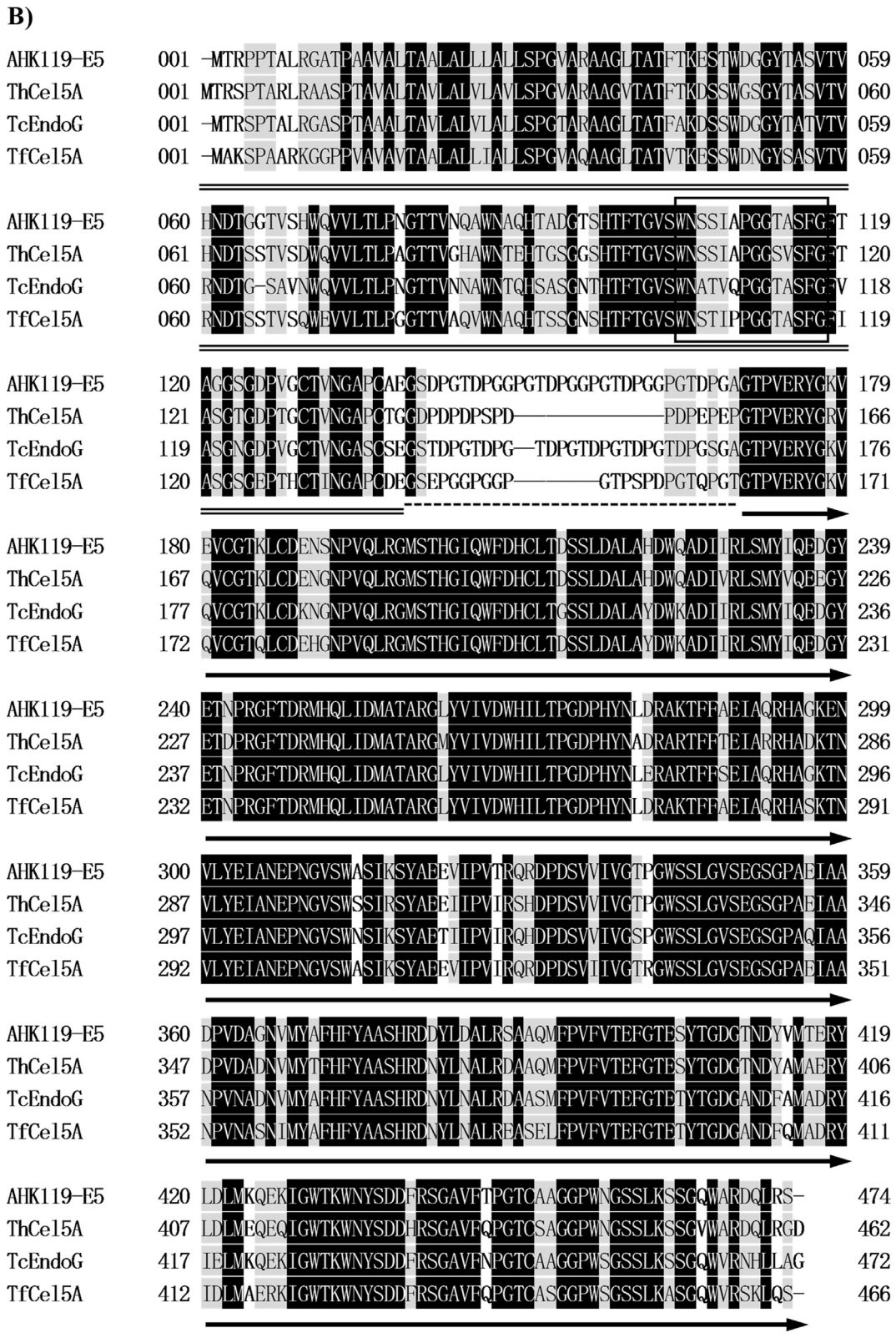

FIG. 2. (*continued*).

fragments were cloned into the pGEM-T easy vector (Promega) and digested with restriction enzymes whose recognition sites were introduced by the primers. Next, the digested fragments were ligated between the BamHI and HindIII sites of pQE30 (Qiagen, Hilden, Germany). The resultant plasmids pQE30-bMs and pQE30-E5 were transformed into *E. coli* BL21(DE3) pLysS. DNA sequencing was performed by a genetic analysis service (SolGent Co., Ltd., Daejeon, Korea). The sequence data were analyzed using BLAST (GenBank, EMBL, and Swiss-Prot databases).

Bioinformatics analysis of cloned sequences DNA and protein sequences of the endoglucanases were analyzed using BLAST (19) and MEGA7 (20). The domain structures of the GH5 endoglucanases were determined using Swiss-Prot, EMBL, NCBI, and CAZy database queries. Phylogenetic tree structure was determined

using the maximum likelihood method based on the JTT matrix-based model (21). Evolutionary analyses were performed using MEGA7 (20). Each protein was named according to the UniProtKB and KEGG databases. The SWISS-MODEL was used for performing 3D modeling of GH5 endoglucanases (22–25). The nucleotide sequence data were deposited into the DDBJ/EMBL/GenBank databases under the accession numbers LC222999 and LC223000 for *AHK119-bMs* and *AHK119-E5*, respectively.

Purification of AHK119-bMs and AHK119-E5 *E. coli* BL21(DE3) pLysS (pQE30-bMs) and (pQE30-E5) were precultured in 4 ml of LB medium containing 50 μg/ml ampicillin and 1% glucose at 37°C overnight. An appropriate amount of the culture was transferred into 100 ml of LB medium containing 50 μg/ml ampicillin and 1% glucose in a 500-ml shake flask and shaken at 150 rpm at 37°C until the OD_{600} reached 0.6. Isopropyl β-D-thiogalactoside was then added to a final concentration of 0.1 mM. The cells were cultured at 27°C for an additional 4 h. Next, the culture was centrifuged, and the cells were collected and suspended in 50 mM sodium phosphate buffer (pH 7.0) containing 300 mM NaCl and 0.1% Tween 20 (buffer A). The cells were disrupted by sonication, and cell debris was removed by centrifugation at 40,000 ×*g* for 20 min at 4°C. The resulting supernatant was applied onto a TALON cobalt resin column (Takara, Kyoto, Japan) equilibrated with buffer A. The column was washed with the same buffer, and the bound protein was eluted with buffer A containing 50 mM imidazole.

Characterization of recombinant GH5 endoglucanases The homogeneities and molecular masses of AHK119-bMs and AHK119-E5 were determined by SDS-PAGE. Blue Star PLUS (Nippon Genetics) was used as a molecular marker protein. Proteins fixed in the gel were visualized with Coomassie Brilliant Blue. The activities of the recombinant GH5 endoglucanases on polysaccharide substrates were determined by measuring the amount of liberated reducing sugars using the ferricyanide method (26). The enzymatic reactions were performed at 50°C for 15 min (mannosidase) and 30 min (cellulase). The mannosidase activity was measured using 2 mg/ml locust bean gum (LBG) (Wako, Tokyo, Japan) or 5% Propol A (Pro-A; Shimizu Chemical, Hiroshima, Japan) as the substrate. The enzymatic reactions were performed using 1.54 μg of AHK119-bMs, 50 mM Tris–HCl buffer (pH 10.0), and a substrate in 0.5-ml reaction volume. Cellulase activity was measured using 2 mg/ml of carboxymethyl cellulose (CMC; Nacalai, Kyoto, Japan) or 20 mg/ml of crystalline cellulose powder (CCel; Nacalai) as the substrate. For AHK119-E5, the enzymatic reactions were performed using 2.15 μg of AHK119-E5, 50 mM Tris–HCl buffer (pH 9.0), and a substrate in 0.5-ml reaction volume. One unit of mannosidase and cellulase activity was defined as the amount of enzyme needed to liberate 1.0 μmole of reducing sugar from the substrate in 1 min and 1 h under optimal conditions, respectively. Exoglucosidase activity was measured using the GOD method with CMC and cellobiose (27). One unit of exoglucosidase activity was defined as the amount of enzyme needed to liberate 1.0 μmole of glucose from the substrate in 1 min under optimal conditions. Specific activity (SA) was calculated as the enzyme units per mg of the purified recombinant enzyme. Thermal stability was determined by the residual activity of the enzymes incubated at 50°C for 0–5 h. The effects of various metal ions and other reagents on the activity of recombinant AHK119-bMs and AHK119-E5 were determined at a final concentration of 1 mM using 0.1% SDS.

RESULTS

Excretion of cellulases in the culture supernatant of *T. alba* AHK119 To confirm the cellulose activity of strain AHK119, it was cultured in LB supplemented with pieces of filter paper. After 5 days, the pieces of filter paper had lost 67.5% of their total weight and had a flaky appearance (Fig. S1) because glucose had been liberated from the filter paper in the culture supernatant. In contrast, control cultures without the inoculum produced no glucose, and no weight loss of the filter paper was observed. Based on these results, the cellulase activity of *T. alba* AHK119 was documented. We attempted to clone cellulase genes from the total DNA of strain AHK119 as described below.

Gene cloning and sequence analysis From the genomic DNA of *T. alba* AHK119, we amplified 1365- and 1425-bp sequences using PCR with primers as described previously. AHK119-bMs and AHK119-E5 were composed of 454 and 474 amino acid residues, and their molecular weights were estimated to be 47.9 and 50.2 kDa, respectively. Nucleotide BLAST searches indicated that *AHK119-bMs* and *AHK119-E5* encoded endo-1,4-β-mannosidase (bMs; 1365 bp) and endo-1,4-β-glucanase (E5; 1425 bp), respectively. MEGA7 phylogenetic analysis based on nucleotide sequences showed that *AHK119-bMs* was most closely related to

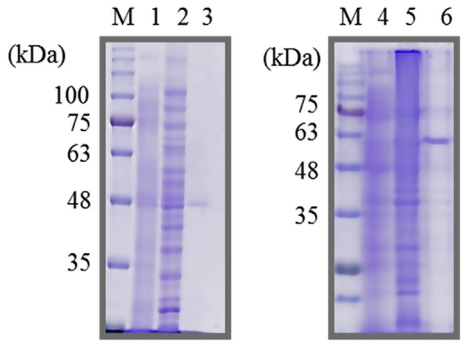

FIG. 3. SDS-PAGE analysis of the expression and purification of recombinant GH5 endoglucanases. Lane M, protein marker. Lanes 1 and 4, the cell extract of *E. coli* BL21(DE3); lane 2, the cell extract of *E. coli* BL21(DE3) expressing pQE30-bMs; lane 3, Co/NTA purified recombinant AHK119-bMs; lane 5, the cell extract of *E. coli* BL21(DE3) expressing pQE30-E5; lane 6, Co/NTA purified recombinant AHK119-E5.

bMs from *T. fusca* YX and endo-β-mannanase (*Man5A*) from *T. fusca* TM51 (*TfMan5A*; Fig. 1A). *AHK119-E5* was most closely related to endo-1,4-β-endoglucanase (*EndoG*) from *T. fusca* YX and cellulase (*Cel5A*) from *T. fusca* (*TfCel5A*; Fig. 1B). *AHK119-E5* was more closely related to *EndoG* and *Cel5A* from thermophilic actinomycetes than to *EndoG* from mesophilic actinomycetes. Protein BLAST searches indicated that the amino acid sequence identities of AHK119-bMs with TfMan5A, Man5A from *T. halotolerans* YIM90462 (ThMan5A), and Man5A from *T. cellulosilytica* TB100 (TcMan5A) were 83%–85% but those with bMs were 59%–63%, except *T. fusca* YX (84%; Fig. S2A). The amino acid sequence identities of AHK119-E5 with EndoG from *T. cellulosilytica* TB100 (TcEndoG), TfCel5A, and Cel5A from *T. halotolerans* YIM90462 (ThCel5A) were 83.4%, 81.9% and 80.6%, respectively. AHK119-E5 was more closely related to ThCel5A than to TfCel5A (Fig. S2B). Fig. 2A and B shows the amino acid sequence alignment of AHK119-bMs and AHK119-E5 with their respective homologous proteins. AHK119-bMs possessed CD on the N-terminal region and CBM2 (28,29) on the C-terminal region, which were connected by the accessory module. On the other hand, AHK119-E5 possessed CBM2 and CD connected by the accessory module, which were in reversed positions compared with AHK119-bMs. These results and the characteristics of CD showed that both AHK119-bMs and AHK119-E5 belonged to the GH5 endoglucanase superfamily. Based on these results, 3D modeling of AHK119-bMs and its homologous Man5As was performed using β-mannanase from *T. fusca* (TfMan; PDB ID: 1BQC) as a template (88% identity with AHK119-bMs; Fig. S3). Additionally, 3D modeling of AHK119-E5, ThCel5A, and TcEndoG was performed using TfCel5A (PDB ID: 2CKS) as a template (87% identity with AHK119-E5; Fig. S4). The Z-scores of the predicted results for 3D modeling were <1, suggesting that the predicted models were of high quality.

Protein expression and purification AHK119-bMs and AHK119-E5 were heterologously expressed in *E. coli* BL21(DE3) as histidine tag-fused proteins at the N-terminal regions and purified from their cell lysates by affinity chromatography on Co/NTA resin. The purity of AHK119-bMs and AHK119-E5 recombinant enzymes was confirmed using SDS-PAGE as single protein bands with apparent molecular weights of approximately 48 and 51 kDa, which were consistent with the theoretical values of 47.9 and 50.2 kDa, respectively (Fig. 3). Using the purified enzymes, AHK119-bMs and AHK119-E5 were characterized as described below.

Characterization of AHK119-bMs The effect of pH on the activity of purified recombinant AHK119-bMs was examined using a variety of buffers ranging from pH 3.0 to 11.2. The activity toward LBG was the highest at pH 9.5, and over 80% of its activity was

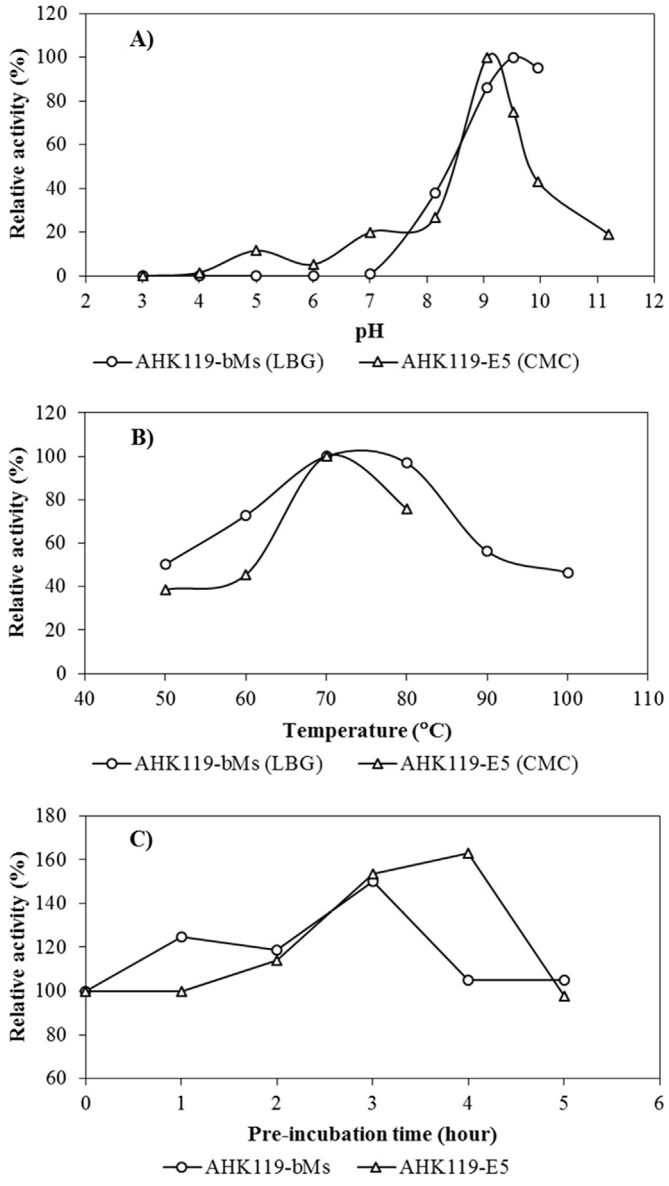

FIG. 4. The effect of pH and temperature on the activity and stability of recombinant GH5 endoglucanases. (A) The effect of pH on the activity. (B) The effect of temperature on the activity. (C) The thermal stability measured by the residual activity after incubation at 50°C for the indicated times. Circles, AHK119-bMs activity toward LBG; triangles, AHK119-E5 activity toward CMC. The data shown in these figures represent three experimental replicates.

retained in the range of pH 9.0 to 10.0 (Fig. 4A). The effect of temperature on the activity of the purified recombinant AHK119-bMs was evaluated in the range of 50–100°C. The activity toward LBG was optimal at 70°C–80°C, and >50% of its activity was retained at 100°C; however, the activity was <80% at 60°C (Fig. 4B). Thermal stability was confirmed by measuring the residual activity after incubation at 50°C for 0–5 h. As shown in Fig. 4C, AHK bMs retained 100% of its activity even after treatment at 50°C for 5 h. Incubation for 1–4 h at 50°C enhanced the enzyme activity by 1.2–1.6-fold compared with no incubation (Fig. 4C). The enzyme was therefore classified as a thermoactive enzyme. AHK119-bMs showed activity toward both LBG (galactomannan) and Pro-A (glucomannan; Table 1), suggesting that the enzyme was a mannanase, as indicated by the phylogenic tree (Fig. S2A). Its SA toward CMC was 1.6-fold higher than that of AHK119-E5 (Tables 1 and 2). As shown in Fig. 5A, the

TABLE 1. Substrate specificities of AHK119-bMs.

Substrate	Enzyme assay	Specific activity (units/mg)
LBG (soluble)	Endo β-1,4-mannosidase activity	1.16[a]
Pro-A (insoluble)	Endo β-1,4-mannosidase activity	0.38[a]
CMC (soluble)	Endo β-1,4-endoglucanase activity	11.9[b]

[a] Calculated by mannosidase unit.
[b] Calculated by cellulase unit.

activity of AHK119-bMs toward soluble LBG was inhibited by the addition of EDTA, SDS, and Cu^{2+} but enhanced by the addition of Zn^{2+} and Mg^{2+} by 2.1- and 1.5-fold, respectively. Remarkably, its activity toward the insoluble substrate Pro-A was enhanced by approximately 5-fold by the addition of Zn^{2+} compared with the control, but Mg^{2+} addition did not show any effect on Pro-A utilization. There is no previous report of a GH5 endo-1,4-β-mannosidase whose activity was enhanced by the addition of Zn^{2+}.

Characterization of AHK119-E5 The effect of pH on the activity of purified recombinant AHK119-E5 was examined at various pH values ranging from pH 3.0 to 11.2. Its activity toward CMC was optimal at pH 9.0; however, at pH <8.0 or >10.0, the activity was decreased to <40% (Fig. 4A). The effect of temperature was evaluated at 50°C–80°C. The activity of AHK119-E5 toward CMC was the highest at 70°C. Although 76% of the activity was retained at 80°C, it was reduced to <50% at 60°C (Fig. 4B). Thermal stability was measured by the residual activity after incubation at 50°C for 0–5 h. Similar to AHK119-bMs, AHK-E5 retained 100% of its activity even after 5 h of incubation, and 2–4 h of incubation enhanced the activity by 1.1–1.6-fold compared with the (Fig. 4C). Thus, AHK119-E5 was a thermoactive and thermostable enzyme. Under optimal assay conditions, AHK119-E5 showed higher endoglucanase activity toward the insoluble substrate CCel than toward the soluble substrate CMC (Table 2). Although slight exoglucosidase activity was also detected with CMC as a substrate, β-glucosidase activity was not detected at all with cellobiose. AHK119-E5 also showed endo-1,4-β-mannosidase activity toward LBG, but it was less than that of AHK119-bMs (Tables 1 and 2). Therefore, AHK119-E5 was considered to be a member of Cel5A. The activity of AHK119-E5 toward soluble and insoluble cellulose was enhanced by Cu^{2+} addition and inhibited by Ca^{2+} addition (Fig. 5B). The activity of TfCel5A was enhanced by Cu^{2+} addition (30), but that of ThCel5A was enhanced by Ca^{2+} addition (31). The roles of Cu^{2+} and Ca^{2+} in the activity of AHK119-E5 would be interesting targets for further research.

DISCUSSION

As summarized in Table 3, the amino acid sequences of CD and CBM2 of AHK119-bMs showed higher identity with those of TfMan5A, ThMan5A, and TcMan5A (32–34) than the overall sequence identity (approximately 84%). Because TfMan consists only of CD, the identity of AHK119 bMs CD and the overall sequence of TfMan were almost identical (97%). Furthermore, 3D modeling of AHK119 bMs and the three Man5As (Fig. S3) showed overall structural similarity in these

TABLE 2. Substrate specificities of AHK119-E5.

Substrate	Enzyme assay	Specific activity (Units/mg)
CMC (soluble)	Endo β-1,4-endoglucanase activity	7.64[a]
CCel (insoluble)	Endo β-1,4-endoglucanase activity	16.8[a]
CMC (soluble)	Exo β-1,4-glucosidase activity	0.10[b]
Cellobiose (soluble)	Exo β-1,4-glucosidase activity	0.00[b]
LBG (soluble)	Endo β-1,4-endoglucanase activity	0.27[c]

[a] Calculated by cellulase unit.
[b] Calculated by glucosidase unit.
[c] Calculated by mannosidase unit.

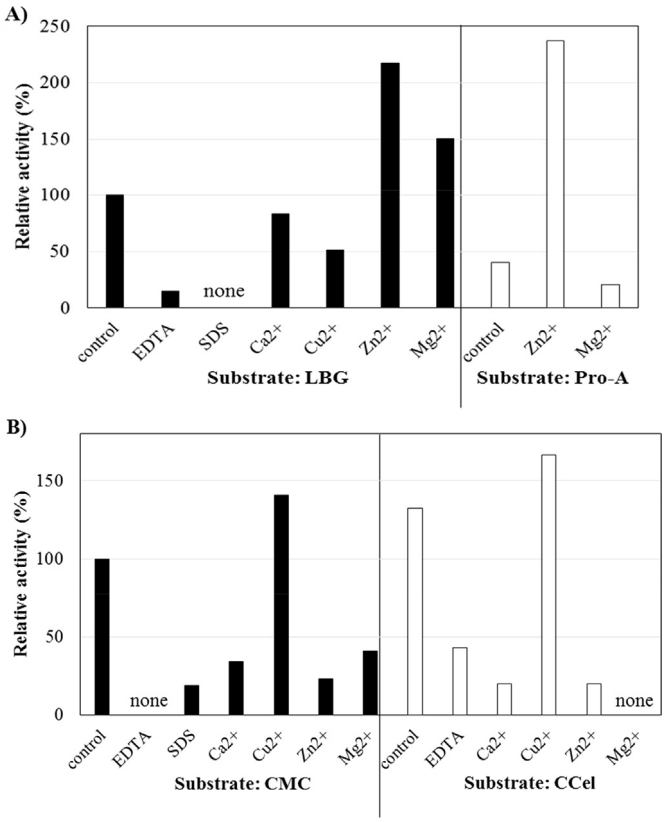

FIG. 5. Effect of metal ions and chemicals on the activity of recombinant GH5 endoglucanases. (A) AHK119-bMs. (B) AHK119-E5. The data shown in these figures represent three experimental replicates.

proteins, and the activity of AHK119 bMs toward LBG and Pro-A strongly supported the fact that AHK119 bMs was a mannanase-like Man5As (Figs. 2A and S2A) rather than a mannosidase (Fig. 1A), as predicted by the phylogenic tree based on the amino acid sequences (Fig. S2A). Notably, the activity of AHK119-bMs was enhanced by Zn^{2+} addition and inhibited by Ca^{2+} addition. TfMan5A and ThMan5A were activated by Ca^{2+} addition, which was in turn related to the resistance to thermal denaturation and halotolerance, respectively (32). Nam et al. (35) revealed that Zn^{2+} is involved in the crystallographic packing and subunit folding of the external region of the TIM-like barrel domain of CelM2, but the activity of CelM2 was not enhanced by Zn^{2+} addition. Although the putative Ca^{2+}-binding site (36,37) was conserved in CD of AHK119 bMs, as found in the three Man5As and TfMan (Ca^{2+} most likely activates TcMan5A, TfMan, TfMan5A, and ThMan5A), the effect of Ca^{2+} addition on the activity of AHK119 bMs was different from that on other enzymes. AHK119 bMs showed appreciable activity toward CMC, although Man5As and TfMan showed no activity toward cellulosic substrates (32,34). However, the surface-exposed tryptophans (W370, W394, and W406) related to substrate binding and cysteines (C360 and C451) related to a disulfide bond were completely conserved in CBM2 of AHK119 bMs and the three Man5As (Fig. 2A). Therefore, the different behaviors of these enzymes in presence of metal ions and cellulosic materials may be dependent on their accessory modules that had low identities with each other.

The identity of the amino acid sequences of AHK119-E5 with TfCel5A and ThCel5A was approximately 82%. X-ray crystallography image of TfCel5A (PDB ID: 2CKS) was used as a template for 3D modeling of AHK119-E5 and ThCel5A (Fig. S3) to demonstrate the high quality of the predicted model. However, optimal pH and temperature, effect of Cu^{2+}, and substrate specificity toward insoluble cellulose (higher than that for soluble cellulose) of AHK119-E5 differed from those of other Cel5As (Table 4). There was a large difference in the optimum pH value between AHK119-E5 and TfCel5A and in the optimum temperature between AHK119-E5 and ThCel5A. The optimum pH and optimum temperature for TfCel5A and ThCel5A also remarkably differed. Cu^{2+} addition enhanced the activity of AHK119-E5 and TfCel5A but inhibited the activity of ThCel5A. On the contrary, Ca^{2+} addition enhanced the activity of ThCel5A, whereas Ca^{2+} addition dramatically inhibited the activity of AHK119-E5 and TfCel5A. TfCel5A and ThCel5A showed activities against Avicel and filter paper, but these activities were considerably lower than those on CMC (31,38,39). The activity of AHK119-E5 toward crystalline cellulose was 2.2-fold higher than that toward CMC (Table 2). The detection of a slight glucose was considered to the result of endoglucanase activity toward CMC-derived cello-oligosaccharide such as cellotriose. The amino acid sequence identities of CBM2 of AHK119-E5 with those of Cel5As and TcEndoG were 73%–77%, with the consensus CBM2a domain signature W-N-[STAGR]-[STDN]-[LIVM]-x(2)-[GST]-x-[GST]-x(2)-[LIVMFT]-[GA] being highly conserved (Fig. 2B). Wilson described that CBM2 of TfCel5A was necessary for the activity toward insoluble substrate, such as cellulose, but not toward soluble substrates (39).

Yan et al. reported that the activity of TfCel5A was enhanced by incubation at 50°C for 10 h and was maintained at 100% after 20 h (30). AHK119-bMs and AHK119-E5 showed similar thermoactivity and thermostability to those of TfCel5A, which indicated the potential for industrial applications.

Taken together, AHK119-bMs and AHK119-E5 had different characteristics from those of other homologous enzymes, although

TABLE 3. Difference between AHK119-bMs and Man5As.

Enzyme	CD/CBM2 identity[a] (%)	Optimal temp. (°C)	Optimal pH	Requirement of metal ion	Activity on cellulosic substrate
AHK119-bMs	–	70–80	9.0–10.0	Zn^{2+} (increase of activity)	Active
TfMan5A	99/84	75	8.0	Ca^{2+} (increase of thermal stability)	None
ThMan5A	99/82	70	7.5	Ca^{2+} (increase of halotolerance)	None
TcMan5A	88/77	70	7.0	None	None
TfMan	97/–	80	7.0–8.0	Unknown	None

[a] Based on the amino acid sequence of AHK119-bMs.

TABLE 4. Difference between AHK119-E5 and Cel5As.

Enzyme	CBM2/CD identity[a] (%)	Optimal temp. (°C)	Optimal pH	Requirement of metal ion	Activity on insoluble substrate (%)[b]
AHK119-E5	–	70–80	9.0	Cu^{2+} (increase of activity)	220
ThCel5A	77/87	50	8.0	Ca^{2+} (increase of activity)	10–90
TfCel5A	73/87	80	5.5	Cu^{2+} (increase of activity)	3

[a] Based on the amino acid sequence of AHK119-E5.
[b] The activity against CMC was taken as 100%.

their functional modules CD and CBM2 shared high identities with those of other homologous enzymes. These findings suggested that factors other than functional modules are strongly related to the characteristics of these endoglucanases. The accessory modules of cellulases are known as catalytically inactive domains, consisting of 10–65 amino acids, and sharing low identities (40). The accessory modules of GH5 endoglucanases also showed low identity with each other. Studies on the linker region of endoglucanases have indicated that the length, flexibility, and rigidity of linker modules affect the kinetic parameters of enzymatic activity (41,42). However, the role of linker region in synergistic action of the catalytic and substrate-binding modules remains to be elucidated, which may help in determining substrate specificities and enzymatic properties.

In this study, we cloned AHK119-bMs and AHK119-E5 from *T. alba* AHK119, which were not identified previously. Both the enzymes possessed a high optimal temperature (80°C) and thermal stability (retained for >5 h at 50°C) and optimal pH in the alkaline range. AHK119-bMs showed high activity toward various cellulosic substrates and AHK119-E5 toward insoluble cellulosic substrates. These characteristics suggest their potential applications in various industrial processes such as bioethanol production, material production, bioprocessing and biorefining.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbiosc.2018.10.015.

References

1. **Malhi, Y.:** Carbon in the atmosphere and terrestrial biosphere in the 21st century, Philos. Trans. A Math. Phys. Eng. Sci., **360**, 2925–2945 (2002).
2. **Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B.:** The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res., **2**(Database issue), D490–D495 (2014).
3. **Wierenga, R. K.:** The TIM-barrel fold: a versatile framework for efficient enzymes, FEBS Lett., **492**, 193–198 (2001).
4. **Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H., and Henrissat, B.:** Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5), BMC Evol. Biol., **12**, 186 (2012).
5. **Tóth, A., Barna, T., Nagy, I., Horváth, B., Nagy, I., Táncsics, A., Kriszt, B., Baka, E., Fekete, C., and Kukolya, J.:** Draft genome sequence of the lignocellulose decomposer *Thermobifida fusca* strain TM51, Genome Announc., **1**, e-00482-13 (2013).
6. **Lykidis, A., Mavromatis, K., Ivanova, N., Anderson, I., Land, M., DiBartolo, G., Martinez, M., Lapidus, A., Lucas, S., Copeland, A., and other 3 authors:** Genome sequence and analysis of the soil cellulolytic actinomycete *Thermobifida fusca* YX, J. Bacteriol., **189**, 2477–2486 (2007).
7. **Posta, K., Béki, E., Wilson, D. B., Kukolya, J., and Hornok, L.:** Cloning, characterization and phylogenetic relationships of cel5B, a new endoglucanase encoding gene from *Thermobifida fusca*, J. Basic Microbiol., **44**, 383–399 (2004).
8. **Béki, E., Nagy, I., Vanderleyden, J., Jäger, S., Kiss, L., Fülöp, L., Hornok, L., and Kukolya, J.:** Cloning and heterologous expression of a β-D-mannosidase (EC 3.2.1.25)-encoding gene from *Thermobifida fusca* TM51, Appl. Environ. Microbiol., **69**, 1944–1952 (2003).
9. **Kukolya, J., Nagy, I., Láday, M., Tóth, E., Oravecz, O., Márialigeti, K., and Hornok, L.:** *Thermobifida cellulosilytica* sp. nov., a novel lignocellulose-decomposing actinomycete, Int. J. Syst. Evol. Microbiol., **52**, 1193–1199 (2002).
10. **Zhang, Z., Wang, Y., and Ruan, J.:** Reclassification of *Thermomonospora* and *Microtetraspora*, Int. J. Syst. Bacteriol., **48**, 411–422 (1998).
11. **Ghangas, G. S. and Wilson, D. B.:** Cloning of the *Thermomonospora fusca* endoglucanase E2 gene in *Streptomyces lividans*: affinity purification and functional domains of the cloned gene product, Appl. Environ. Microbiol., **54**, 2521–2526 (1988).
12. **Larsson, A. M., Bergfors, T., Dultz, E., Irwin, D. C., Roos, A., Driguez, H., Wilson, D. B., and Jones, T. A.:** Crystal structure of *Thermobifida fusca* endoglucanase Cel6A in complex with substrate and inhibitor: the role of tyrosine Y73 in substrate ring distortion, Biochemistry, **44**, 12915–12922 (2005).
13. **McGrath, C. E. and Wilson, D. B.:** Characterization of a *Thermobifida fusca* β-1,3-glucanase (Lam81A) with a potential role in plant biomass degradation, Biochemistry, **45**, 14094–14100 (2006).
14. **van Wyk, N., den Haan, R., and van Zyl, W. H.:** Heterologous co-production of *Thermobifida fusca* Cel9A with other cellulases in *Saccharomyces cerevisiae*, Appl. Microbiol. Biotechnol., **87**, 1813–1820 (2010).
15. **Caspi, J., Barak, Y., Haimovitz, R., Gilary, H., Irwin, D. C., Lamed, R., Wilson, D. B., and Bayer, E. A.:** *Thermobifida fusca* exoglucanase Cel6B is incompatible with the cellulosomal mode in contrast to endoglucanase Cel6A, Syst. Synth. Biol., **4**, 193–201 (2010).
16. **Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A.:** Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*, Nat. Struct. Biol., **4**, 810–818 (1997).
17. **Blanco, J., Coque, J. J., Velasco, J., and Martín, J. F.:** Cloning, expression in *Streptomyces lividans* and biochemical characterization of a thermostable endo-β-1,4-xylanase of *Thermomonospora alba* ULJB1 with cellulose-binding ability, Appl. Microbiol. Biotechnol., **48**, 208–217 (1997).
18. **Hu, X., Osaki, S., Hayashi, M., Kaku, M., Katuen, S., Kobayashi, H., and Kawai, F.:** Degradation of terephthalate-containing polyester by thermophilic actinomycetes and Bacillus species derived from composts, J. Polym. Environ., **16**, 103–108 (2008).
19. **Madden, T. L., Tatusov, R. L., and Zhang, J.:** Applications of network BLAST server, Methods Enzymol., **266**, 131–141 (1996).
20. **Kumar, S., Stecher, G., and Tamura, K.:** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol., **33**, 1870–1874 (2016).
21. **Jones, D. T., Taylor, W. R., and Thornton, J. M.:** The rapid generation of mutation data matrices from protein sequences, Comput. Appl. Biosci., **8**, 275–282 (1992).
22. **Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., and Schwede, T.:** SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res., **42**, 252–258 (2014).
23. **Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T.:** The SWISS-MODEL Repository and associated resources, Nucleic Acids Res., **37**, 387–392 (2009).
24. **Arnold, K., Bordoli, L., Kopp, J., and Schwede, T.:** The SWISS-MODEL Workspace: a web-based environment for protein structure homology modelling, Bioinformatics, **22**, 195–201 (2006).
25. **Guex, N., Peitsch, M. C., and Schwede, T.:** Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective, Electrophoresis, **30**, 162–173 (2009).
26. **Hulme, A. C. and Narain, R.:** The ferricyanide method for the determination of reducing sugars, A modification of the Hagedorn-Jensen-Hanes technique, Biochem. J., **25**, 1051–1061 (1931).
27. **Miwa, I., Okudo, J., Maeda, K., and Okuda, G.:** Mutarotase effect on colorimetric determination of blood glucose with D-glucose oxidase, Clin. Chim. Acta, **37**, 538–540 (1972).
28. **Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., and Warren, R. A.:** Domains in microbial beta-1,4-glycanases: sequence conservation, function, and enzyme families, Microbiol. Rev., **55**, 303–315 (1991).
29. **Nagy, T., Simpson, P., Williamson, M. P., Hazlewood, G. P., Gilbert, H. J., and Orosz, L.:** All three surface tryptophans in Type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands, FEBS Lett., **429**, 312–316 (1998).
30. **Yan, P., Su, L., Chen, J., and Wu, J.:** Heterologous expression and biochemical characterization of an endo-β-1,4-glucanase from *Thermobifida fusca*, Biotechnol. Appl. Biochem., **60**, 348–355 (2013).
31. **Zhang, F., Zhang, X. M., Yin, Y. R., and Li, W. J.:** Cloning, expression and characterization of a novel GH5 exo/endoglucanase of *Thermobifida halotolerans* YIM 90462^T by genome mining, J. Biosci. Bioeng., **120**, 644–649 (2015).
32. **Tóth, Á., Barna, T., Szabó, E., Elek, R., Hubert, Á., Nagy, I., Kriszt, B., Táncsics, A., and Kukolya, J.:** Cloning, expression and biochemical characterization of endomannanases from *Thermobifida* species isolated from different niches, PLoS One, **11**, e0155769 (2016).
33. **Hilge, M., Gloor, S. M., Rypniewski, W., Sauer, O., Heightman, T. D., Zimmermann, W., Winterhalter, K., and Piontek, K.:** High-resolution native and complex structures of thermostable β-mannanase from *Thermomonospora fusca* — substrate specificity in glycosyl hydrolase family 5, Structure, **6**, 1433–1444 (1998).
34. **Hilge, M., Gloor, S., Winterhalter, K., Zimmermann, W., and Piontek, K.:** Crystallization and preliminary crystallographic analysis of two beta-mannanase isoforms from *Thermomonospora fusca* KW3, Acta Crystallogr. D Biol. Crystallogr., **52**, 1224–1225 (1996).
35. **Nam, K. H., Lee, W. H., Rhee, K. H., and Hwang, K. Y.:** Structural characterization of the bifunctional glucanase-xylanase CelM2 reveals the metal effect and substrate-binding moiety, Biochem. Biophys. Res. Commun., **391**, 1726–1730 (2010).
36. **Kumagai, Y., Usuki, H., Yamamoto, Y., Yamasato, A., Arima, J., Mukaihara, T., and Hatanaka, T.:** Characterization of calcium ion sensitive region for β-mannanase from *Streptomyces thermolilacinus*, Biochim. Biophys. Acta, **1814**, 1127–1133 (2011).
37. **Kumagai, Y., Kawakami, K., Mukaihara, T., Kimura, M., and Hatanaka, T.:** The structural analysis and the role of calcium binding site for thermal stability in mannanase, Biochimie, **94**, 2783–2790 (2012).

38. **Wilso, D. B.:** Cellulases of *Thermomonospora fusca*, Methods Enzymol., **160**, 314–323 (1988).
39. **Wilson, D. B.:** Studies of *Thermobifida fusca* plant cell wall degrading enzymes, Chem. Rec., **4**, 72–82 (2004).
40. **Sammond, D. W., Payne, C. M., Brunecky, R., Himmel, M. E., Crowley, M. F., and Beckham, G. T.:** Cellulase linkers are optimized based on domain type and function: insights from sequence analysis, biophysical measurements, and molecular simulation, PLoS One, **7**, e48615 (2012).
41. **Srisodsuk, M., Reinikainen, T., Penttilä, M., and Teeri, T. T.:** Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose, J. Biol. Chem., **268**, 20756–20761 (1993).
42. **Ruiz, D. M., Turowski, V. R., and Murakami, M. T.:** Effects of the linker region on the structure and function of modular GH5 cellulases, Sci. Rep., **6**, 28504 (2016).