Characterization of a Novel Glycosyl Hydrolase 12 Family Endoglucanase from *Talaromyces pinophilus* Y117

Peng Li [iD],[#] Cheng Zhang,[#] Junhui Nie, Dawei Xiong, Shuaiwen Zhang, Siyuan Yue, Jing Zeng, and Lin Yuan [iD],*

Endoglucanase plays a vital role in lignocellulose degradation, yet the functional diversity of glycosyl hydrolase 12 (GH12) endoglucanases remains underexplored. In this study, TpCel12a, a novel GH12 endoglucanase from *Talaromyces pinophilus* Y117 was identified and characterized. It exhibited optimal activity at pH 4.0 and 40 °C with a preference for xyloglucan. Size exclusion chromatography revealed a monomeric state of TpCel12a. Despite classification into the promiscuous subfamily I, TpCel12a showed strict specificity for xyloglucan and carboxymethyl cellulose, yielding XXXG-type oligosaccharides and glucose/cellobiose, respectively. A loop 3 insertion mutant, designated as TpCel12a(+TQA), enhanced substrate affinity (5-fold lower K_m) but reduced activity, suggesting a trade-off between binding and catalysis. Surprisingly, TpCel12a inhibited the hydrolytic efficiency of native *T. pinophilus* crude cellulase on corn cob powder, coinciding with its undetectable secretion under microcrystalline cellulose induction. This study highlights TpCel12a's unique biochemical properties and its unexpected antagonistic role in native cellulase systems, offering insights for engineering fungal enzyme cocktails.

INTRODUCTION

Lignocellulosic biomass, such as forestry residues (Cotana *et al*. 2014) and agricultural waste (Yuan *et al*. 2018), represents the most abundant renewable organic resource on earth (Saldarriaga-Hernández *et al*. 2020). Lignocellulosic biomass is mainly composed of cellulose, hemicellulose (including xyloglucan and xylan), and lignin, with cellulose being the most abundant, representing about 40% to 60% in weight (Sharma *et al*. 2019). Cellulose is a linear polymer with D-glucose as its monomeric unit linked by β-1,4-glycosidic bonds. In fungi, typical cellulose degradation is thought to be accomplished by an enzymatic cocktail, collectively referred to as cellulase, in a synergistic manner. Cellulase consists of three major types of enzymes: cellobiohydrolases (CBHs, CBH I: EC 3.2.1.91 and CBH II: EC 3.2.1.176), which cleave cellulose from the reducing and non-reducing ends, respectively, resulting in the release of cellobiose; endo-β-1,4-glucanases (EC 3.2.1.4), which randomly cleave the internal bonds of cellulose; and β-glucosidases

(EC 3.2.1.21), which hydrolyze cellobiose, resulting in the release of glucose monomers (Ohmiya *et al*. 1997). In addition to cellulase, lytic polysaccharide monooxygenases (LPMOs) from auxiliary activity families 9 and 10 (AA9 and AA10) have been demonstrated to boost cellulase activity by deconstructing cellulose (Villares *et al*. 2017; Guo *et al*. 2022).

It has been shown that cellulases with endoglucanases as the major component have diverse industrial applications, such as in cotton and fruit juice processing (Sharma *et al*. 2017), brewing and oil extraction (Ejaz *et al*. 2021). Importantly, endoglucanases play a vital role in the bioconversion of cellulosic biomass into fermentable sugars (Periyasamy *et al*. 2023), which can be further used to produce bioethanol. Therefore, there has been a growing demand for endoglucanases in recent years.

Currently, endoglucanases are mainly produced by cellulolytic microorganisms, especially fungi, such as *Fomitopsis* species (Patel and Shah 2021), *Aspergillus species* (Dobrev and Zhekova 2012), *Fusarium* species (de Almeida *et al*. 2013), *Gloeophyllum* species (Cohen *et al*. 2005), and *Trichoderma* species (Zheng *et al*. 2024). Among these, *Trichoderma reesei* has been the predominant cellulase producer for industrial applications over the past decades. Endoglucanases are grouped into various glycoside hydrolase families based on amino acid sequence similarity and mode of action in the CAZy database (Drula *et al*. 2022), including GH5, GH6, GH7, GH8, GH9, GH10, GH12, GH26, GH44, GH45, GH48, GH51, GH74, and GH124. Distinct from other endoglucanases, all characterized GH12 family members exclusively possess a single catalytic domain without CBM, resulting in their relatively low molecular mass. The relatively low molecular mass characteristic of GH12 family endoglucanases is thought to promote their diffusion into plant cell wall matrices, thereby improving cellulose depolymerization (Cohen *et al*. 2005; Miotto *et al*. 2014).

Numerous GH12 enzymes have been identified and functionally characterized, exhibiting diverse enzymatic activities, including endoglucanase (Wang *et al*. 2017), xyloglucanase (Matsuzawa *et al*. 2020), and β-1,3-1,4-glucanase activities (Miotto *et al*. 2014). The functional diversity underscored the versatility and potential applications of GH12 enzymes in various biotechnology processes.

Talaromyces pinophilus (formerly classified as *Penicillium pinophilum*) is a filamentous fungal species known for its robust ability to produce biomass-degrading enzymes (Yamanobe *et al*. 1987; Visser *et al*. 2013; Xian *et al*. 2015). Due to this capacity, it has emerged as a promising alternative microbial platform for efficient cellulase production and biomass degradation. However, GH12 enzymes derived from *Talaromyces species* have been rarely documented in existing studies. In previous work, the authors successfully isolated a strain, named Y117, of *T. pinophilus* with high cellulase production. Here, a novel GH12 family endoglucanase, designated TpCel12a, was identified in *T. pinophilus* strain Y117 genome.

The recombinant TpCel12a was successfully expressed in its native host by integrating the TpCel12a coding sequence under the *cbh1* promoter. The enzymatic action of TpCel12a on xyloglucan and CMC were investigated. In addition, the synergistic action of TpCel12a coupled with the crude cellulase was evaluated on the degradation of a natural lignocellulosic substrate, corn cob powder.

EXPERIMENTAL

Materials

Strains and plasmids used in the study are listed in Appendix I. *Escherichia coli* TOP10 strain, employed as the host for vector construction and propagation, was cultured in LB medium supplemented with appropriate antibiotics when required. *T. pinophilus* strain Y117 was cultured in the medium according to the protocol described by Fang *et al*. (2008). The preculture medium for *T. pinophilus* strain Y117 contained 40 g/L microcrystalline cellulose, 24 g/L KH_2PO_4, 5 g/L $(NH_4)_2SO_4$, 1.2 g/L $MgSO_4 \cdot 7H_2O$, 10 mg/L $ZnSO_4 \cdot 7H_2O$, 10 mg/L $MnSO_4 \cdot 6H_2O$, 10 mg/L $CuSO_4 \cdot 7H_2O$, and 2 g/L urea. The cellulase production and protein overexpression medium consisted of 50 g/L microcrystalline cellulose, 24 g/L KH_2PO_4, 5 g/L $(NH_4)_2SO_4$, 1.2 g/L $MgSO_4 \cdot 7H_2O$, 10 mg/L $ZnSO_4 \cdot 7H_2O$, 10 mg/L $MnSO_4 \cdot 6H_2O$, 10 mg/L $CuSO_4 \cdot 7H_2O$, and 4 g/L urea. The pH was adjusted to 4.0 with H_2SO_4 and KOH before sterilization. A 100×trace element stock solution (1.0 g/L $ZnSO_4 \cdot 7H_2O$, 1.0 g/L $MnSO_4 \cdot 6H_2O$, 1.0 g/L $CuSO_4 \cdot 7H_2O$) was prepared and sterilized individually. The urea was dissolved in distilled water, sterilized, and filtered through a 0.22 µm filter membrane. The other components of the medium were dissolved in distilled water and sterilized at 121 °C for 20 min. The pUC57 plasmid served as the backbone for constructing the TpCel12a and its mutant expression vector. Carboxymethyl cellulose (CMC, medium viscosity) was obtained from Sigma-Aldrich (St. Louis, MO, USA). Tamarind xyloglucan, β-1,3-1,4-glucan, D-galactosyl-D-mannan, β-D-mannan, xylan, arabinogalactan, pectin, and standards (glucose, cellobiose, cellotriose, cellotetraose, cellopentaose, and cellohexaose) were acquired from Psaitong Biotechnology Co., Ltd (Beijing, China). Primers used in this study were listed in Appendix II.

Methods

Plasmid construction

The plasmid pBIP-TpCel12a for overexpressing TpCel12a was constructed as follows. The terminator region of the cbh1 gene from *Talaromyces rugulosus* strain W13939 (GenBank accession: GCA_013368755, Chromosome I: 839,585–839,934 nt) was synthesized, and the corresponding homologous sequence was inserted using primers TRcbh1-ter-f/r (Appendix II). The promoter region of the *cbh1* gene (2 kb) from *T. pinophilus* strain Y117, serving as a homologous arm, along with the coding sequence of *tpCel12a*, were amplified from *T. pinophilus* Y117 genomic DNA using primer pairs TPcbh1-prom-f/r and tpcel12a-f/r (Appendix II), respectively. Subsequently, splicing by overlap extension PCR (SOE-PCR) was performed with primers TPcbh1-prom-f/TRcbh1-ter-r (Appendix II) to assemble the three fragments into a single construct, designated TPcbh1prom-tpCel12a-TRcbh1ter. This fragment was then cloned between the *Eco*RI and *Sac*I sites of the pUC57 vector using a seamless cloning kit (Vazyme, China), yielding the intermediate plasmid pUC-TPcbh1prom-tpCel12a-TRcbh1ter. Additionally, the promoter region of the housekeeping gene pyruvate kinase from *T. rugulosus* W13939 (GCA_013368755, Chromosome I: 2,479,882–2,480,381 nt), and the cbh1 terminator from *T. stipitatus* ATCC 10500 (GCA_000003125, scaffold scf_110550729515: 901,690–902,039 nt) were synthesized, with homologous sequences inserted using primers TRpk-prom-f/r and TScbh1-ter-f/r (Appendix II), respectively.

The coding sequence of the hygromycin B phosphotransferase gene (*hph*), conferring hygromycin resistance, was amplified from the plasmid pLKO.1-Hygro. SOE-PCR was conducted with primers TRpk-prom-f/TScbh1-ter-r (Appendix II) to fuse these three fragments into a single unit, TRpkprom-hyg-TScbh1ter, which was subsequently cloned between the *Sma*I and *Pst*I sites of the intermediate plasmid pUC-TPcbh1prom-tpCel12a-TRcbh1ter using a seamless cloning kit. Finally, the expression plasmid pBIP-TpCel12a was generated.

Inverse PCR with site-directed primers TpCel12a-inTQA-f/r (Appendix II) was performed to introduce a TQA insertion into loop 3 of TpCel12A. All constructs were verified by DNA sequencing.

Construction of TpCel12a overexpression mutant

TpCel12a overexpression mutant was constructed as illustrated in Appendix IV. Briefly, the plasmid pBIP-TpCel12a was linearized by *Not*I and then introduced into *T. pinophilus* Y117 by protoplast transformation. The protoplast preparation and transformation were performed as previously described (Zhao *et al*. 2016). Hygromycin-resistant transformants were selected and expression of TpCel12a was confirmed by western blot.

Sequence analysis

The potential signal peptide and N-glycosylation sites of TpCel12a were predicted using SignalP 6.0 (Teufel *et al*. 2022) and NetNGlyc 1.0 (Gupta and Brunak 2002) servers, respectively. The model structure of TpCel12a was constructed by Alphafold 3.0 (Abramson *et al*. 2024). Multiple sequence alignment was performed using the Clustal Omega online server (Sievers and Higgins 2014), and the results were visualized using ESPript 3.0 server (Gouet *et al*. 2003). Phylogenetic analysis of TpCel12a was conducted using MEGA X software with the maximum likelihood method, employing 500 bootstrap replicates to assess node support (Kumar *et al*. 2018).

Expression and purification of TpCel12a in T. pinophilus strain Y117

The TpCel12a overexpression mutant was inoculated into 50 mL of fresh preculture medium and incubated for 5 days to prepare seed cultures. For TpCel12a production, 10% (v/v) of the seed culture was transferred into 100 mL of production medium in 1 L Erlenmeyer flasks and cultivated at 28 °C with agitation at 250 rpm for 7 days. The culture supernatant was collected following centrifugation at 12,000 × g and 4 °C for 30 min. The supernatant was applied to a Ni-NTA column (Cytiva, Sweden) pre-equilibrated with binding buffer (20 mM Tris, 300 mM NaCl, 20 mM imidazole, pH 8.0), and the proteins were eluted with 20 mL elution buffer (20 mM Tris-HCl, 300 mM imidazole and 300 mM NaCl, pH 8.0) after washing. The eluted protein was further desalted by ultrafiltration with buffer (20 mM Tris-HCl, 10% glycerol and 150 mM NaCl, pH 8.0). The purified protein was then visualized by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The concentration was determined using a Bradford Protein Assay Kit (Beyotime, China). Finally, the oligomeric state of TpCel12a was analyzed by size-exclusion chromatography (SEC) using a low molecular weight (LMW, molecular weight of standard proteins range from 6.5 to 75.0 kDa) calibration kit (Cytiva, Sweden).

Identification and sequencing of TpCel12a N-terminus by liquid chromatography-tandem mass spectrometry (LC-MS/MS)

LC-MS/MS was applied to determine the N-terminal sequence of TpCel12a. Generally, the purified TpCel12a were reduced with 10 mM DTT, alkylated with iodoacetaminde in the dark at room temperature, and digested with endoproteinase Glu-C using a standard protocol. Detailed procedures are described in the Appendix III. Prior to analysis, the peptides were reconstituted in 10 μL of 0.1% (v/v) formic acid. Then LC-MS/MS was performed to analyze the digests. Detailed procedures were provided in the Appendix III.

TpCel12a activity assay and substrate specificity

Endoglucanase activity of TpCel12a was determined by measuring the amount of reducing sugars produced from xyloglucan or CMC with the 3,5-dinitrosalicylic acid (DNS) method (Miller 1959). The reaction mixture was prepared by mixing 10 μL purified TpCelo12a of appropriately diluted enzyme solution with 90 μL of 10 mg/mL xyloglucan or CMC dissolved in 50 mM sodium citrate buffer (pH 4.0) at 50 ℃ for 30 min. About 0.1 mL DNS was added to stop the hydrolysis reaction. The tubes were placed in the boiling water for 5 min, and then they were immediately cooled to room temperature. After diluting the reaction mixtures with 0.8 mL dd-H_2O, the absorbance was measured at 540 nm. One unit (U) activity was defined as the amount of enzyme that produced 1.0 μmol of glucose equivalent per minute under standard assay conditions. TpCel12a content was determined by the Bradford method using bovine serum albumin (BSA) as a standard.

Substrate specificity was examined in assays using a range of substrates. Substrates were prepared as 10 mg/mL in citrate buffer (pH 4.0). The tested substrates were xyloglucan, CMC, β-1,3-1,4-glucan, D-galactosyl-D-mannan, β-D-mannan, xylan, arabinogalactan, and pectin. The reaction mixtures were prepared as described above. The maximum activity of substrate was defined as 100%, TpCel12a activity towards other substrates was calculated according to the maximum activity.

Biochemical characterization of TpCel12a and its mutant

The optimum conditions for activity of TpCel12a were determined by assaying the enzyme activity in a wide range of pH and temperatures with xyloglucan as substrate. In order to determine the optimal pH for activity, the following buffers (each 50 mM) were used: glycine-HCl buffer for pH 2.0 and 3.0, sodium citrate buffer for pH 3.0 and 4.0, sodium acetate buffer for pH 4.0, 5.0 and 6.0, sodium phosphate buffer for pH 6.0, 7.0 and 8.0. The reaction mixture was incubated at 50 ℃ for 30 min. And then the reducing sugars were quantified with the described above method. pH stability of TpCel12a was assessed by measuring the residual activity after incubation in the described above buffers for 1 h at room temperature without substrate.

To investigate the optimal temperature for TpCel12a activity, the citrate buffer with pH 4.0 was used. The reaction mixtures were incubated at different temperatures (30 ℃ to 90 ℃) for 30 min. TpCel12a thermostability was analyzed by detecting the residual activity after incubation at 30 to 80 ℃ for 1 h without substrate.

Kinetic parameters of TpCel12a and its mutant

K_m and V_{max} values of TpCel12a and its mutant were determined in sodium acetate buffer (pH 4.0, 50 mM) containing 1.0 to 10 mg/mL xyloglucan at 40 ℃ (TpCel12a) or 50

°C (TpCel12a(+TQA)) for 5 min. The set of experiments was carried out three times, and each individual experiment included triplicate tests. The kinetic parameters were calculated based on the Lineweaver-Burk or Michaelis-Menten plotting using Graphpad.

Effects of additives on the activity of TpCel12a

The influence of metal ions (Na^+, K^+, NH_4^+, Cu^{2+}, Fe^{3+}, Ca^{2+}, Zn^{2+} and Mn^{2+}) and chemical reagents (SDS, DMSO, EDTA, glycerol, Triton X-100, isopropanol and methanol) on TpCel12a activity were analyzed. Different reagents with final concentration 5 mM or 10% (v/v) were added into the reaction mixtures. And the reaction was carried out in optimal activity conditions for 30 min. The reducing sugars were determined using the described above method. Reaction mixture without any additives were defined as control. All assays were performed in triplicate.

Analysis of xyloglucan and CMC hydrolysis products catalyzed by TpCel12a

Matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy (MALDI-TOF MS) analysis was performed to determine the final hydrolysis products of xyloglucan or CMC catalyzed by TpCel12a. Detailed procedure could be found in Appendix I. For determination of the final hydrolysis products of CMC or corncob powder catalyzed by TpCel12a, high performance liquid chromatography (HPLC) was carried out as described in the Supplementary Materials. Glucose, cellobiose, cellotriose, cellotetraose, cellopentaose, and cellohexaose were used as standards.

Synergistic action between TpCel12a and crude cellulase

The crude cellulolytic enzyme was prepared from *T. pinophilus* strain Y117 under microcrystalline cellulose induction conditions, in which TpCel12a was not detected by secretome proteomics (unpublished). The synergistic effect of TpCel12a and the crude cellulase was evaluated by comparing the degradation efficiency of their mixture to that of either component alone at pH 4 and at 40 °C. The degradation reactions were performed for 48 h with orbital shaking (240 rpm). TpCel12a (20 U) and crude cellulose (0.4 mg) were co-incubated in reaction mixtures containing 0.2 g of untreated corn cob powder in a final volume of 3 mL sodium acetate buffer (50 mM, pH 4) at 40°C for 48 h with orbital shaking (240 rpm). The reducing sugars released in the supernatant were quantified using the DNS method, with glucose as the standard, as described previously. All assays were conducted in triplicate.

RESULTS AND DISCUSSION

Identification and Sequence Analysis of Endoglucanase TpCel12A of *T. pinophilus* Strain Y117

The draft genome of *T. pinophilus* strain Y117 (Accession No.: JBKFBR-000000000) was analyzed for CAZy family assignments using the CAZy database, identifying a gene encoding a putative endoglucanase. The mRNA sequence of this gene was deposited in the GenBank database under accession number PP913528. Based on the deduced amino acid sequence, the endoglucanase was classified into the glycoside hydrolase family 12 and designated as TpCel12a. The TpCel12a gene comprised 834 bp, including two introns of 59 and 64 nucleotides, respectively. The open reading frame was

711 bp in length and encoded 236 amino acid residues. A putative Sec/SPI type signal peptide (MKLTFLLNLAVAASA) was predicted in TpCel12a. The theoretical molecular weight (MW) and isoelectric point (pI) of mature TpCel12a were 23.96 kDa and pH 5.04, respectively. Two potential N-glycosylation sequons (N43-W44-S45 and N145-G146-S147) was predicted within TpCel12a (Fig. 1).
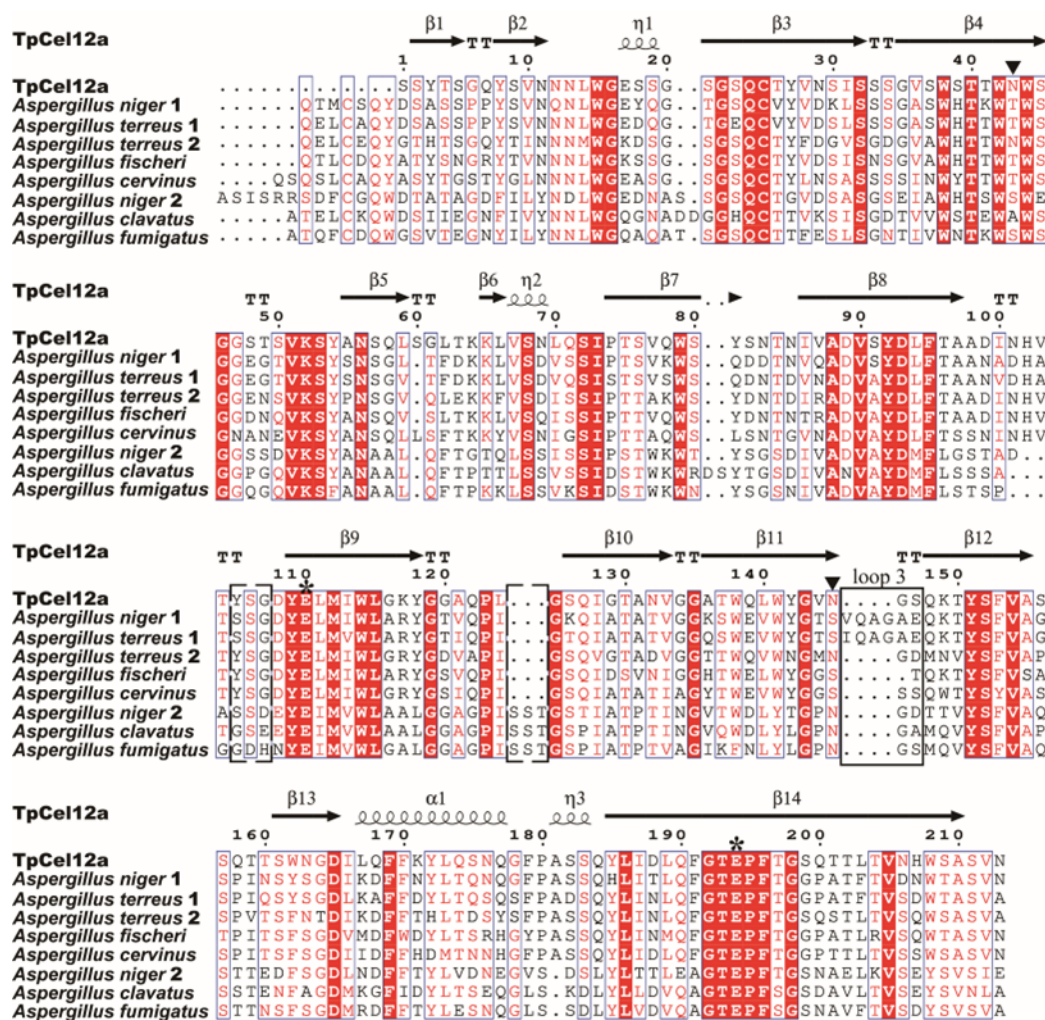

Fig. 1. Multi sequence alignment among TpCel12a and the known GH12 endoglucanases. Amino acid sequences of TpCel12a and other known GH12 endoglucanases including Q0CLB5_ASPTN (1) and AtEglD (Q0C8U0_ASPTN, 2) of *Aspergillus terreus*, Bgh12A (A0A3G2C3I4_9EURO) of *Aspergillus cervinus*, AclaXegA (XGEA_ASPCL) of *Aspergillus clavatus*, NfGH12A (A0A1L6CE30 _9EURO) of *Aspergillus fischeri*, A0A100IKG6_ASPNG (1) and A0A100I2E1_ASPNG (2) of *Aspergillus niger* and XGEA_ASPFU of *Aspergillus fumigatus* were aligned using Clustal and visualized with ESPript. The secondary structural features of TpCel12a were presented as the model structure created by Alphafold 3. The asterisk indicates the conserved catalytic sites E110 and E194. The solid triangle indicates the potential glycosylation site. The dashed rectangular box outlines the typical features YSG and SST motif of xyloglucanase and promiscuous endoglucanase. The solid box marks the loop3 region of GH12 endoglucanases.

Multiple sequence alignments revealed that TpCel12A shared considerable sequence identity with reported GH12 endoglucanases: 69.23% with Bgh12A, 67% with AtEglD, and 43% with AclaXegA, indicating that these enzymes may originate from a common ancestor and have retained their core functional regions during long-term evolution. Like other GH12 enzymes, TpCel12a adopted a β-sandwich structure. Additionally, amino acid residues (E110 and E194) responsible for the catalytic activity of GH12 endoglucanases were conserved in the β9 and β14 regions of TpCel12a, respectively (Fig. 1), illustrating that TpCel12a follows the typical catalysis mechanism of the GH12 family.

Overexpression and Purification of TpCel12A in *T. pinophilus* strain Y117

Given the robust cellulase biosynthesis and secretory capacity of *T. pinophilus* strain Y117, TpCel12a was overexpressed with its native signal peptide in the native host under the control of the *cbh1* promoter (Appendix IV). The TpCel12a overexpression mutant OEtpCel12a was constructed as described in the Methods section. TpCel12a expression was induced using microcrystalline cellulose as the sole carbon source. After purification by Ni-NTA affinity chromatography,
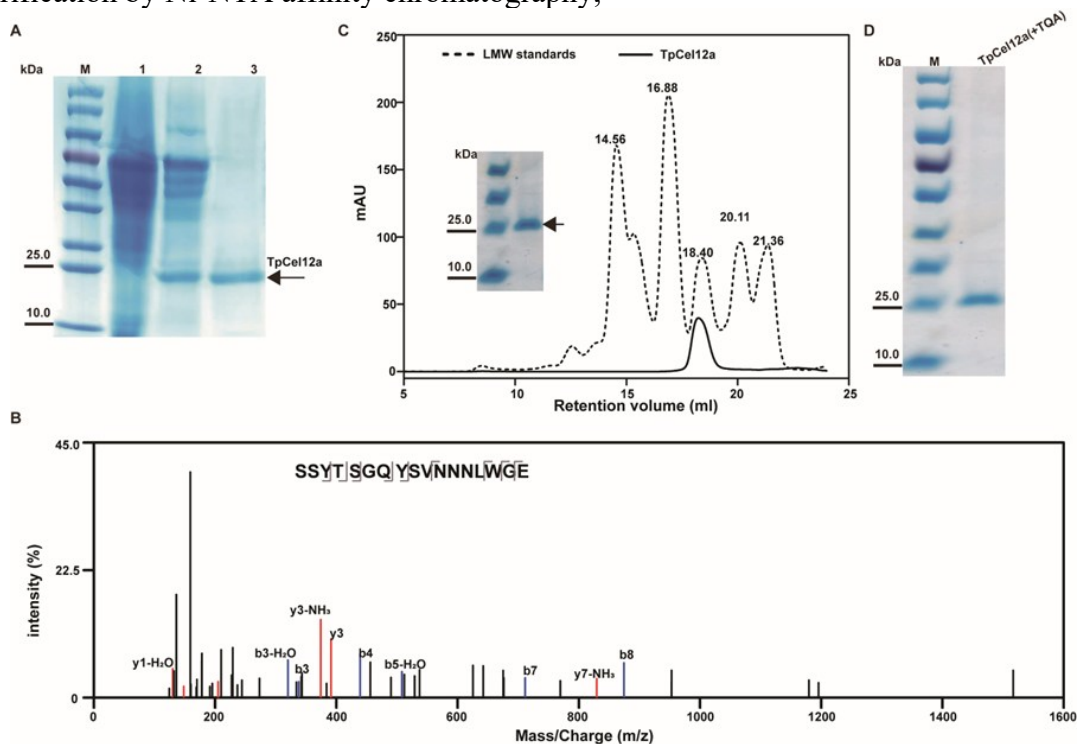

Fig. 2. Over-expression and purification of TpCel12a and the mutant TpCel12a(+TQA) with a C-terminal His-tag in *T. pinophilus* strain Y117. A) Recombinant TpCel12a was over-expressed in *T. pinophilus* Y117 and purified by Ni-column affinity chromatography. Lane M: protein marker, lane 1: total proteins of *T. pinophilus* without TpCel12a expression plasmid, lane 2: total proteins of *T. pinophilus* with TpCel12a expression plasmid, lane 3: purified recombinant TpCel12a. B) N-terminal sequence of mature TpCel12a was determined by do novo N- terminal sequencing. C) Size-exclusion chromatography was carried out to detect the oligomeric state of recombinant TpCel12a. Dashed lines represent the elution volumes of protein standards (provided in LMW (low molecular weight) calibration kit, molecular weight of LMW standard proteins ranges from 6.5 to 75.0 kDa), with the numbers above the peaks indicating the retention volume. D) Overexpression and purification of the mutant TpCel12a(+TQA).

SDS-PAGE analysis revealed a single protein band corresponding to the theoretical molecular weight of TpCel12a (Fig. 2A). The successful expression of TpCel12A in *T. pinophilus* strain Y117 suggests that this strain has potential to be engineered into a biosynthetic platform for enzyme production from cost-effective lignocellulosic biomass, pending further optimization of key elements such as promoters and signal peptides.

De novo N-terminal sequencing further identified the cleavage site of the signal peptide, yielding the N-terminal sequence SSYTSGQYSVNNNLWGE (Fig. 2B). This result confirmed the presence of a 23-amino acid signal peptide, which was notably longer than the bioinformatically predicted 15-amino acid signal peptide. To determine the oligomeric state of recombinant TpCel12a, we performed size-exclusion chromatography. Recombinant TpCel12a was eluted at 18.24 mL (Fig. 2C). Based on the calibration curve (Appendix V), the apparent MW of recombinant TpCel12a was determined to be 24.44 kDa, which is consistent with the theoretical MW (24.1 kDa) of the mature TpCel12a protein fused with a C-terminal 8×His-tag. This result suggests that recombinant TpCel12a exists as a monomer in solution and TpCel12a was not glycosylation modified at the putative glycosylation site.
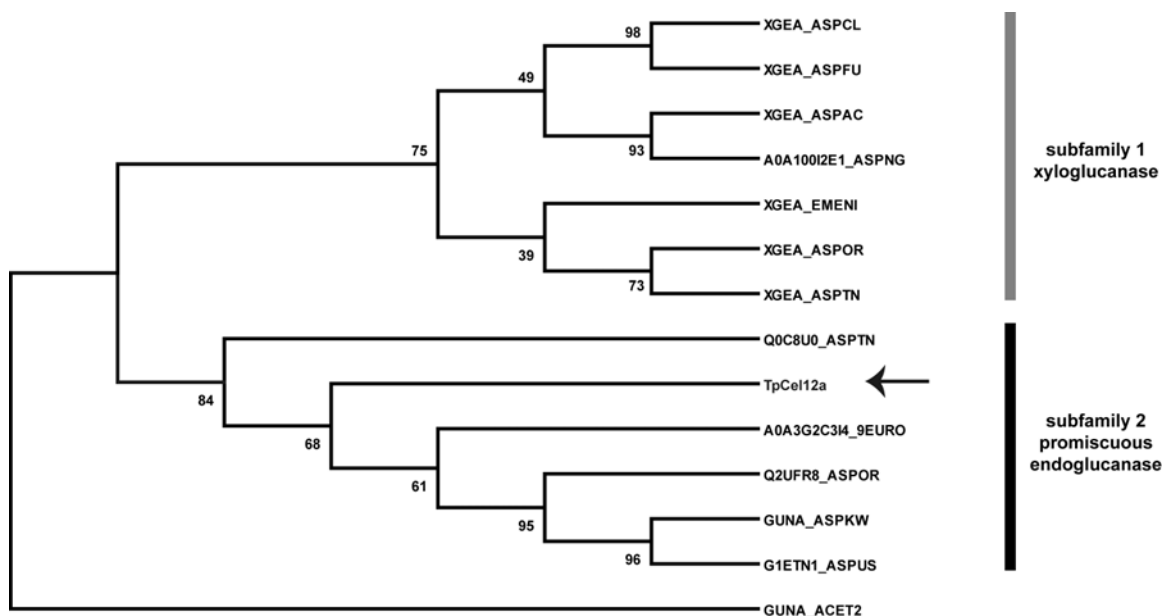

Fig. 3. Phylogenetic analysis of GH12 family endoglucanases. The phylogenetic tree was constructed using the maximum likelihood method with MEGA X. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (500) are shown next to the branches. The amino acid sequences of GH12 family endoglucanases were retrieved from UniProtKB database (Lussi *et al.* 2023), including XGEA_ASPCL (*Aspergillus clavatus* ATCC1007), XGEA_ASPFU (*Aspergillus fumigatus* (ATCC MYA-4609)), XGEA_ASPAC (*Aspergillus aculeatus* KSM510), A0A100I2E1_ASPNG (Aspergillus nigerAn76), XGEA_EMENI (*Emericella nidulans* ATCC38163), XGEA_ASPOR (*Aspergillus oryzae* ATCC42149), XGEA_ASPTN (*Aspergillus terreus* NIH2624), Q0C8U0_ASPTN (Aspergillus terreus NIH2624), A0A3G2C3I4_9EURO (*Aspergillus cervinus* VKPM F-612), Q2UFR8_ASPOR (*Aspergillus oryzae* ATCC42149), GUNA_ASPKW (*Aspergillus kawachii* NBRC4308), G1ETN1_ASPUS (*Aspergillus usamii* ATCC11364) and GUNA_ACET2 (*Acetivibrio thermocellus* ATCC27405) of GH8 family, which was used as outgroup. The black arrow indicates the target sequence TpCel12a. Two distinct subfamilies representing xyloglucanase and promiscuous endoglucanase were marked by gray and black lines, respectively.

Substrate Specificity of Recombinant TpCel12a

Previous studies revealed that GH12 enzymes from Eurotiomycetes were divided into two distinct subgroups, subgroup I and II: subfamily I exhibits general endoglucanase activity, while subfamily II demonstrates xyloglucan-specific activity (Rykov *et al*. 2019). Phylogenetic analysis revealed that TpCel12a clusters within subfamily I (Fig. 3). In addition, typical sequence features, presence of the YSG motif and absence of the SST motif (Damásio et al. 2014), were observed in TpCel12a (Fig. 1). The analysis suggests that TpCel12a might function as a promiscuous endoglucanase.

To investigate the substrate specificity, the recombinant TpCel12a was performed to hydrolyze various isolated polysaccharides. The results demonstrated that TpCel12a can hydrolyze both CMC and xyloglucan. And TpCel12a exhibited the maximum activity toward xyloglucan, with approximately 30% of that activity observed on CMC. However, TpCel12a showed almost no activity towards insoluble microcrystalline cellulose (MCC). This phenomenon can be explained by the dense, highly aggregated crystalline structure of MCC, which physically prevents TpCel12a from accessing the β-1,4-glycosidic bonds in the glucan backbone and thus inhibits catalytic activity. It was therefore hypothesized that TpCel12a would display intermediate activity on amorphous cellulose, which offers greater accessibility than MCC but less than fully soluble polymers. Furthermore, to precisely delineate the substrate specificity and cellulolytic activity of TpCel12a, future work will include assays using highly pure cellulosic substrates with minimal hemicellulose content, such as cotton fibers or bleached kraft pulps.

Interestingly, TpCel12a showed no observable hydrolysis activity against β-1,3-1,4-glucan substrates, which serve as common substrates for many GH12 subfamily I enzymes. Compared with GH12 enzymes capable of binding β-1,3-1,4-glucan (Ma *et al*. 2021), TpCel12a's substrate-binding pocket may be unable to accommodate the conformational changes of the polysaccharide chain caused by the β-1,3 glycosidic bond. Additionally, TpCel12a exhibited no detectable hydrolytic activity against other polysaccharides, such as D-galactosyl-D-mannan, β-D-mannan, xylan, arabinogalactan, and pectin. This substrate preference not only conforms to the functional characteristics of GH12 Subfamily I, but also exhibits a certain degree of specificity, providing theoretical support for its application in specific polysaccharide degradation (such as the degradation of mixed substrates of xyloglucan and cellulose).

Biochemical Characterization and Enzymatic Kinetics of TpCel12a

TpCel12a showed the highest hydrolytic activity toward xyloglucan. Therefore, the optimal conditions for activity were investigated with xyloglucan as the substrate. The optimal temperature and pH were determined to be 40 °C and pH 4, respectively (Figs. 4A and 4B). Comparative analysis of buffer systems revealed that TpCel12a displayed higher enzymatic activity in acetate buffer than in citrate buffer at equivalent pH values (Fig. 4B), suggesting that in industrial applications, the catalytic efficiency of enzymes can be further improved by optimizing the type of buffers in the reaction system. The enzyme exhibited stability across a broad pH range (3.0 to 8.0), retaining >90% of its initial activity after 1 h incubation at room temperature (Fig. 4D). The pH stability of TpCel12a is comparable to that of Xgh12B (Rykov *et al*. 2019), but superior to that of TaCel12 from *Trichoderma asperellum* ND-1 (Zheng *et al*. 2024). The wide pH stability might allow TpCel12a adapt to the pH ranges of various industrial systems reducing the loss of enzyme activity caused by pH fluctuations. Thermal stability assays showed complete retention of activity

following 1.0 h treatments at 30 °C and 40 °C. However, the enzyme was completely inactivated at temperatures ⩾ 50 °C (Fig. 4C), revealing the limited thermostability of recombinant TpCel12A. The absence of glycosylation modification in TpCel12a may affect its thermal stability (Anbarasan *et al.* 2010). Therefore, introducing new glycosylation sites or exposing potential ones would improve the thermal stability.
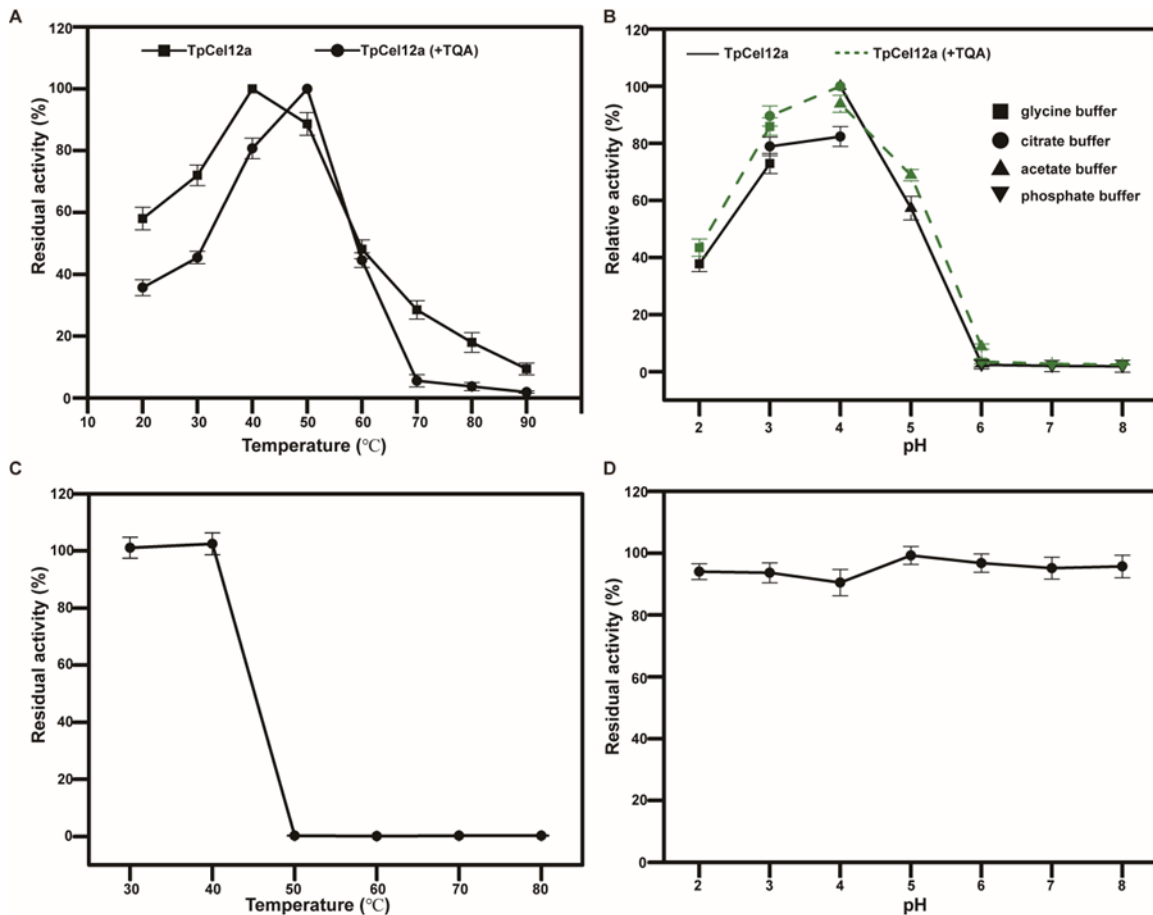

Fig. 4. The optimal conditions of recombinant TpCel12a and its mutant TpCel12a(+TQA). A) Optimal temperature of TpCel12a and its mutant TpCel12a(+TQA). Solid squares (■) and circles (●) represent TpCel12a and TpCel12a(+TQA), respectively. B) Optimal pH of TpCel12a and its mutant TpCel12a(+TQA). Glycine (■), citrate (●), acetate (▲), and phosphate (▼) buffers are indicated by different symbols. The black solid line represents TpCel12a, while the green dashed line denotes TpCel12a(+TQA). Relative activities are expressed as percentages of the maximum activity (set as 100%). C) Thermostability profile of TpCel12a. D) pH stability profile of TpCel12a. Relative activities are expressed as percentages of the untreated control (set as 100%). Data represent mean values ± SD (n = 3).

To investigate the effects of metal ions and chemical reagents on TpCel12A activity, reactions were supplemented with either 5 mM metal ions or 10% (v/v) chemical reagents. As shown in Fig. 5A, most tested metal ions showed little or no effect on enzymatic activity, with the exception of Mn^{2+}, which reduced activity to below 60% of the control. In contrast, chemical reagents exhibited more pronounced inhibitory effects (Fig. 5B). SDS and EDTA almost completely abolished TpCel12A activity, while glycerol and Triton X-100 preserved approximately 80% of the original activity.
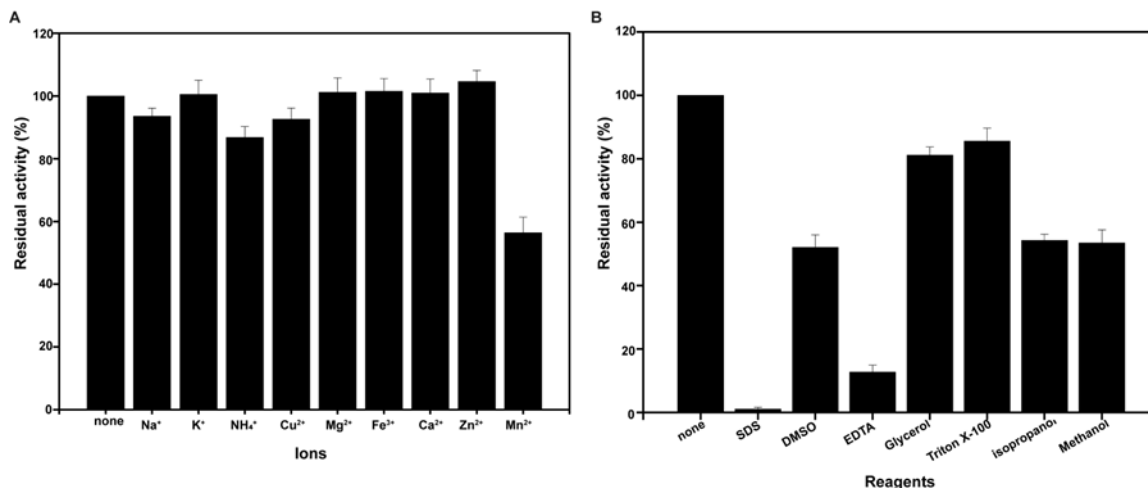
Fig. 5. Effects of metal ions and chemical reagents on the activity of TpCel12a. A) Influence of metal ions on enzymatic activity. B) Influence of chemical reagents on enzymatic activity. Relative activities are expressed as percentages of the untreated control (set as 100%). Data represent mean values ± SD (n = 3).

Moderate inhibition was observed with DMSO, isopropanol or methanol, each maintaining about 50% residual activity. The kinetic parameters of TpCel12a acting on xyloglucan were measured at the optimal condition and determined using Lineweaver-Burk plots. The enzyme exhibited a K_m of 27.13±1.54 mg/mL and a V_{max} of 559.02±5.32 μmol/(mg·min). The K_m value was much higher than those of previously reported GH12 endoglucanase, such as 2.81±0.17 mg/mL of Xgh12B (Rykov *et al*. 2019) and 0.174±0.008 mg/mL of Xeg12A (Matsuzawa *et al*. 2020), suggesting a weaker substrate affinity. However, the V_{max} of TpCel12a was much higher that of Xeg12A (270±3 μmol/(mg·min)) (Matsuzawa *et al*. 2020). The combination of high K_m and V_{max} indicated TpCel12a as a potential candidate for xyloglucan degradation in biomass refining processes.

Enzymatic Hydrolysis Patterns of TpCel12a on Xyloglucan and CMC

To analyze the hydrolysis pattern of TpCel12a on xyloglucan degradation, MALDI-TOF MS was used to examine the final hydrolysis products of xyloglucan treated with TpCel12a. The results demonstrated that after 6 h of enzymatic hydrolysis, as monitored by reducing sugar quantification to confirm reaction completion, four oligosaccharides including XXXG, XLXG, XXLG, and XLLG were generated (Figs. 6A and 6B). This cleavage pattern aligns with the previously reported mode of xyloglucan degradation by GH12 family members, but differs from that of GH5 family enzymes, which typically produce shorter oligosaccharide units (Matsuzawa *et al*. 2020). Similarly, the end products of CMC hydrolysis by TpCel12a were analyzed and identified glucose, cellobiose, and cellotriose as the major products, with trace amounts of cellotetraose and cellopentaose also detected (Appendix VI). The authors further investigated the hydrolytic products of TpCel12a on CMC using HPLC. By comparing retention times with known standards, it was determined that TpCel12a mainly generates glucose and cellobiose as the final hydrolysis products (Fig. 6C), which differs from the product profile of processive endoglucanases (Schiano-di-Cola *et al*. 2020), implying the weak processivity of TpCel12a.
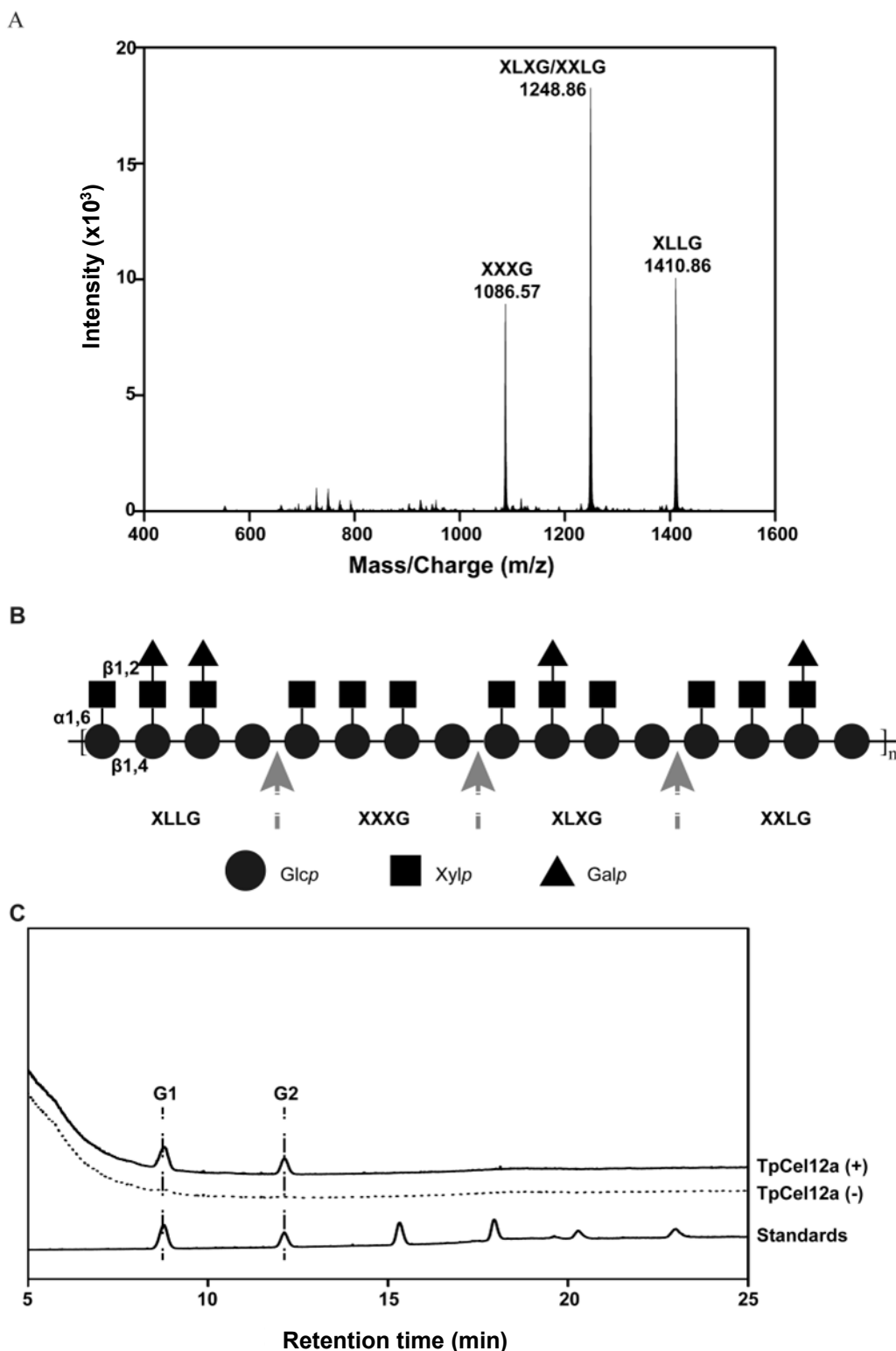
Fig. 6. The final hydrolysis products of xyloglucan and CMC catalyzed by TpCel12a. A) MALDI-TOF MS determination of final products of xyloglucan catalyzed by TpCel12a. The XXXG-type oligosaccharides with their molecular weight were indicated above their corresponding peaks. B) The hydrolysis mode of TpCel12a on xyloglucan. C) HPLC determination of final products of CMC catalyzed by TpCel12a. G1 and G2 indicate glucose and cellobiose, respectively.

Mutation of Loop3 Significantly Increased the TpCel12a's Affinity for Xyloglucan

The loop3 region of GH12 endoglucanases has been shown to play a critical role in modulating catalytic efficiency. Insertion mutation of loop3 was found to enhance the catalysis efficiency of NfEG12A (Yang *et al*. 2017). Multiple sequence alignment revealed that TpCel12a also possesses a short loop3 as NfEG12A. To examine whether extending loop3 could enhance the catalytic efficiency of TpCel12a, a three-amino-acid sequence (TQA) was inserted between residues N154 and G155, generating the mutant TpCel12a(+TQA). This variant was overexpressed and purified using the same protocol as wild-type TpCel12a (Fig. 2D). The mutation shifted the optimal temperature from 40 to 50 °C (Fig. 4A), but there was no discernible effect on thermostability (not shown). The optimal pH and pH stability of TpCel12a and TpCel12a(+TQA) were nearly identical, except that TpCel12a(+TQA) exhibited higher activity in citrate buffer (pH 4.0) compared to acetate buffer (pH 4.0) (Fig. 4B). Kinetic analysis based on Michaelis-Menten plots revealed a K_m of 3.94 ± 0.15 mg/mL and a V_{max} of 198.58 ± 4.32 μmol/(min·mg). The affinity for xyloglucan increased 5-fold, as indicated by the decrease in K_m. However, the enzymatic activity was reduced, indicating a trade-off relationship between affinity and catalytic activity.

Synergistic Action in Saccharification of Untreated Corn Cob Powder between TpCel12a and Crude Cellulase

The synergistic action of exoglucanase, endoglucanase, β-glucosidase, and even auxiliary enzyme such as LPMO in cellulase is essential for the efficient degradation of lignocellulose into soluble sugars. Previous studies have shown that certain GH12 endoglucanases can enhance hydrolysis efficiency when combined with crude cellulase (Narra *et al*. 2014). Therefore, to examine the potential synergistic effect between TpCel12a and crude cellulase, TpCel12a was combined with the crude cellulase from its native host *T. pinophilus* Y117 in a reaction mixture containing untreated corn cob powder as substrate. Hydrolysis was performed at 40 °C and pH 4.0.
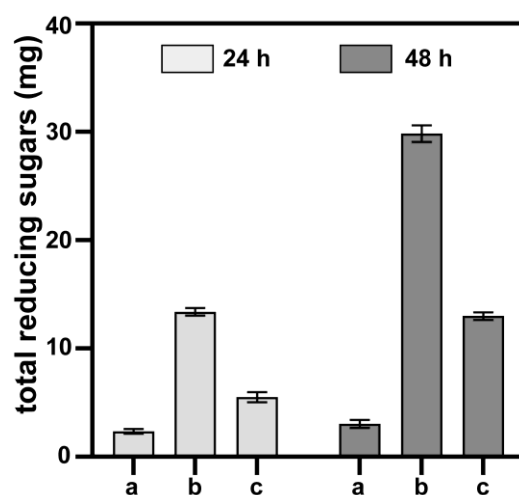

Fig. 7. No synergistic effect was observed between TpCel12a and crude cellulose from its native host *T. pinophilus* Y117: a: TpCel12a only; b: crude cellulase only; and c: combination of TpCel12a and crude cellulase.

While TpCel12a alone showed limited activity on untreated corn cob powder compared to the crude cellulase alone (Fig. 7), the combination of TpCel12a with crude cellulase unexpectedly resulted in significantly reduced hydrolysis efficiency (Fig. 7). This suggests that TpCel12a not only failed to synergize with the crude cellulase, but that it also might inhibit the hydrolytic activity of the *T. pinophilus* Y117 crude cellulase. The inhibition might be explained by the fact that cellobiose—a major product of TpCel12a—is a well-known inhibitor of exoglucanases and endoglucanases in crude cellulase mixtures (Yue et al. 2004). Moreover, this finding may explain why LC-MS analysis failed to detect secretory expression of TpCel12a when induced by microcrystalline cellulose (not shown). Further experiments are needed to determine whether TpCel12a shows synergistic activity with other enzymes including xylanases, mannanases, or cellulases of different origins.

CONCLUSIONS

1. A novel GH12 endoglucanase TpCel12a from *T. pinophilus* strain Y117 was identified and characterized.

2. Despite the fact that TpCel12a belongs to promiscuous phylogenetic subfamily I, TpCel12a exhibited substrate specificity for xyloglucan/CMC, and no detected activity towards β-1,3-1,4-glucan and other polysaccharides.

3. TpCel12a hydrolyzed xyloglucan into XXXG-type oligosaccharides and CMC into glucose/cellobiose, respectively.

4. Through insertion mutagenesis in loop 3 TpCel12a(+TQA), a 5-fold reduction was achieved in the kinetic parameter K_m (27.13 mg/mL to 3.94 mg/mL), significantly improving substrate binding affinity, albeit at the cost of reduced enzymatic activity.

5. Unexpectedly, TpCel12a showed no synergistic effect with crude cellulase during untreated corn cob powder saccharification, instead, it exhibited inhibitory effects on crude cellulase activity.

ACKNOWLEDGMENTS

REFERENCES CITED

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., *et al*. (2024). "Accurate structure

prediction of biomolecular interactions with AlphaFold 3," *Nature* 630(8016), 493-500. https://doi.org/10.1038/s41586-024-07487-w

Anbarasan, S., Jänis, J., Paloheimo, M., Laitaoja, M., Vuolanto, M., Karimäki, J., Vainiotalo, P., Leisola, M., and Turunen, O. (2010). "Effect of glycosylation and additional domains on the thermostability of a family 10 xylanase produced by *Thermopolyspora flexuosa*," *Appl. Environ. Microbiol*. 76(1), 356-360. https://doi.org/10.1128/AEM.00357-09

Cohen, R., Suzuki, M. R., and Hammel, K. E. (2005). "Processive endoglucanase active in crystalline cellulose hydrolysis by the brown rot basidiomycete *Gloeophyllum trabeum*," *Appl. Environ. Microbiol.* 71(5), 2412-2417. https://doi.org/10.1128/aem.71.5.2412-2417.2005

Cotana, F., Cavalaglio, G., Gelosia, M., Nicolini, A., Coccia, V., and Petrozzi, A. (2014). "Production of bioethanol in a second generation prototype from pine wood chips," *Energy Procedia* 45, 42-51. https://doi.org/10.1016/j.egypro.2014.01.006

Damásio, A. R., Rubio, M. V., Oliveira, L. C., Segato, F., Dias, B. A., Citadini, A. P., Paixão, D. A., and Squina, F. M. (2014). "Understanding the function of conserved variations in the catalytic loops of fungal glycoside hydrolase family 12," *Biotechnol. Bioeng.* 111(8), 1494-1505. https://doi.org/10.1002/bit.25209

de Almeida, M. N., Falkoski, D. L., Guimarães, V. M., Ramos, H. J., Visser, E. M., Maitan-Alfenas, G. P., and de Rezende, S. T. (2013). "Characteristics of free endoglucanase and glycosidases multienzyme complex from *Fusarium verticillioides*," *Bioresour. Technol.* 143, 413-22. https://doi.org/10.1016/j.biortech.2013.06.021

Dobrev, G. T., and Zhekova, B. Y. (2012). "Biosynthesis, purification and characterization of endoglucanase from a xylanase producing strain *Aspergillus niger* B03," *Braz. J. Microbiol.* 43(1), 70-77. https://doi.org/10.1590/S1517-83822012000100008

Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). "The carbohydrate-active enzyme database: Functions and literature," *Nucleic Acids Res.* 50(D1), D571-D577. https://doi.org/10.1093/nar/gkab1045

Ejaz, U., Sohail, M., and Ghanemi, A. (2021). "Cellulases: From bioactivity to a variety of industrial applications," *Biomimetics* 6(3), article 44. https://doi.org/10.3390/biomimetics6030044

Fang, X., Yano, S., Inoue, H., and Sawayama, S. (2008). "Lactose enhances cellulase production by the filamentous fungus *Acremonium cellulolyticus*," *J. Biosci. Bioeng*. 106(2), 115-120. https://doi.org/10.1263/jbb.106.115

Gouet, P., Robert, X., and Courcelle, E. (2003). "ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins," *Nucleic Acids Res.* 31(13), 3320-3323. https://doi.org/10.1093/nar/gkg556

Guo, X., An, Y., Jiang, L., Zhang, J., Lu, F., and Liu, F. (2022). "The discovery and enzymatic characterization of a novel AA10 LPMO from *Bacillus amyloliquefaciens* with dual substrate specificity," *Int. J. Biol. Macromol.* 203, 457-465. https://doi.org/10.1016/j.ijbiomac.2022.01.110

Gupta, R., and Brunak, S. (2002). "Prediction of glycosylation across the human proteome and the correlation to protein function," *Pac. Symp. Biocomput.* 310-322.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). "MEGA X: Molecular evolutionary genetics analysis across computing platforms," *Mol. Biol. Evol.* 35(6), 1547-1549. https://doi.org/10.1093/molbev/msy096

Lussi, Y. C., Magrane, M., Martin, M. J., and Orchard, S. (2023). "Searching and navigating UniProt databases," *Curr. Protoc.* 3(3), article e700. https://doi.org/10.1002/cpz1.700

Ma, J., Li, Y., Han, S., Jiang, Z., Yan, Q., and Yang, S. (2021). "Structural and biochemical insights into the substrate-binding mechanism of a glycoside hydrolase family 12 β-1,3-1,4-glucanase from *Chaetomium* sp.," *J. Struct. Biol*. 213(3), article 107774. https://doi.org/10.1016/j.jsb.2021.107774

Matsuzawa, T., Kameyama, A., Nakamichi, Y., and Yaoi, K. (2020). "Identification and characterization of two xyloglucan-specific endo-1,4-glucanases in *Aspergillus oryzae*," *Appl. Microbiol. Biotechnol.* 104(20), 8761-8773. https://doi.org/10.1007/s00253-020-10883-7

Miller, G. L. (1959). "Use of dinitrosalicylic acid reagent for determination of reducing sugar," *Analytical Chemistry* 31(3), 426-428. https://doi.org/10.1021/ac60147a030

Miotto, L. S., de Rezende, C. A., Bernardes, A., Serpa, V. I., Tsang, A., and Polikarpov, I. (2014). "The characterization of the endoglucanase Cel12A from *Gloeophyllum trabeum* reveals an enzyme highly active on β-glucan," *PLoS One* 9(9), article e108393. https://doi.org/10.1371/journal.pone.0108393

Narra, M., Dixit, G., Divecha, J., Kumar, K., Madamwar, D., and Shah, A. R. (2014). "Production, purification and characterization of a novel GH 12 family endoglucanase from *Aspergillus terreus* and its application in enzymatic degradation of delignified rice straw," *International Biodeterioration and Biodegradation* 88, 150-161. https://doi.org/10.1016/j.ibiod.2013.12.016

Ohmiya, K., Sakka, K., Karita, S., and Kimura, T. (1997). "Structure of cellulases and their applications," *Biotechnol. Genet. Eng. Rev.* 14, 365-414. https://doi.org/10.1080/02648725.1997.10647949

Patel, A., and Shah, A. (2021). "Purification and characterization of novel, thermostable and non-processive GH5 family endoglucanase from *Fomitopsis meliae* CFA 2," *Int. J. Biol. Macromol.* 182, 1161-1169. https://doi.org/10.1016/j.ijbiomac.2021.04.110

Periyasamy, S., Isabel, J. B., Kavitha, S., Karthik, V., Mohamed, B. A., Gizaw, D. G., Sivashanmugam, P., and Aminabhavi, T. M. (2023). "Recent advances in consolidated bioprocessing for conversion of lignocellulosic biomass into bioethanol–A review," *Chemical Engineering Journal* 453, article 139783. https://doi.org/10.1016/j.cej.2022.139783

Rykov, S. V., Kornberger, P., Herlet, J., Tsurin, N. V., Zorov, I. N., Zverlov, V. V., Liebl, W., Schwarz, W. H., Yarotsky, S. V., and Berezina, O. V. (2019). "Novel endo-(1,4)-β-glucanase Bgh12A and xyloglucanase Xgh12B from *Aspergillus cervinus* belong to GH12 subgroup I and II, respectively," *Appl. Microbiol. Biotechnol.* 103(18), 7553-7566. https://doi.org/10.1007/s00253-019-10006-x

Saldarriaga-Hernández, S., Velasco-Ayala, C., Leal-Isla Flores, P., de Jesús Rostro-Alanis, M., Parra-Saldivar, R., Iqbal, H. M. N., and Carrillo-Nieves, D. (2020). "Biotransformation of lignocellulosic biomass into industrially relevant products with the aid of fungi-derived lignocellulolytic enzymes," *Int. J. Biol. Macromol.* 161, 1099-1116. https://doi.org/10.1016/j.ijbiomac.2020.06.047

Schiano-di-Cola, C., Kołaczkowski, B., Sørensen, T. H., Christensen, S. J., Cavaleiro, A. M., Windahl, M. S., Borch, K., Morth, J. P., and Westh, P. (2020). "Structural and biochemical characterization of a family 7 highly thermostable endoglucanase from the fungus *Rasamsonia emersonii*," *Febs J.* 287(12), 2577-2596. https://doi.org/10.1111/febs.15151

Sharma, H. K., Xu, C., and Qin, W. (2019). "Biological pretreatment of lignocellulosic biomass for biofuels and bioproducts: An overview," *Waste and Biomass Valorization* 10, 235-251. https://doi.org/10.1007/s12649-017-0059-y

Sharma, H. P., Patel, H., and Sugandha. (2017). "Enzymatic added extraction and clarification of fruit juices – A review," *Crit. Rev. Food Sci. Nutr.* 57(6), 1215-1227. https://doi.org/10.1080/10408398.2014.977434

Sievers, F., and Higgins, D. G. (2014). "Clustal omega," *Curr. Protoc. Bioinformatics* 48, 3.13.1-3.13.16. https://doi.org/10.1002/0471250953.bi0313s48

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2022). "SignalP 6.0 predicts all five types of signal peptides using protein language models," *Nat Biotechnol* 40(7), 1023-1025. https://doi.org/10.1038/s41587-021-01156-3

Villares, A., Moreau, C., Bennati-Granier, C., Garajova, S., Foucat, L., Falourd, X., Saake, B., Berrin, J. G., and Cathala, B. (2017). "Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure," *Sci. Rep.* 7, article 40262. https://doi.org/10.1038/srep40262

Visser, E. M., Falkoski, D. L., de Almeida, M. N., Maitan-Alfenas, G. P., and Guimarães, V. M. (2013). "Production and application of an enzyme blend from *Chrysoporthe cubensis* and *Penicillium pinophilum* with potential for hydrolysis of sugarcane bagasse," *Bioresour. Technol.* 144, 587-94. https://doi.org/10.1016/j.biortech.2013.07.015

Wang, Z., Hu, Y., Long, L., and Ding, S. (2017). "Characterization of a GH12 endoglucanase from *Volvariella volvacea* exhibiting broad substrate specificity and potential synergy with crude cellulase," *BioResources* 12(4), 9437-9451. https://doi.org/10.15376/biores.12.4.9437-9451

Xian, L., Wang, F., Luo, X., Feng, Y. L., and Feng, J. X. (2015). "Purification and characterization of a highly efficient calcium-independent α-amylase from *Talaromyces pinophilus* 1-95," *PLoS One* 10(3), article e0121531. https://doi.org/10.1371/journal.pone.0121531

Yamanobe, T., Mitsuishi, Y., and Takasaki, Y. (1987). "Isolation of a cellulolytic enzyme producing microorganism, culture conditions and some properties of the enzymes," *Agricultural and Biological Chemistry* 51(1), 65-74. https://doi.org/10.1080/00021369.1987.10867998

Yang, H., Shi, P., Liu, Y., Xia, W., Wang, X., Cao, H., Ma, R., Luo, H., Bai, Y., and Yao, B. (2017). "Loop 3 of fungal endoglucanases of glycoside hydrolase family 12 modulates catalytic efficiency," *Appl. Environ. Microbiol.* 83(6). https://doi.org/10.1128/AEM.03123-16

Yuan, Z., Wen, Y., and Li, G. (2018). "Production of bioethanol and value added compounds from wheat straw through combined alkaline/alkaline-peroxide pretreatment," *Bioresour. Technol.* 259, 228-236. https://doi.org/10.1016/j.biortech.2018.03.044

Yue, Z., Bin, W., Baixu, Y., and Peiji, G. (2004). "Mechanism of cellobiose inhibition in cellulose hydrolysis by cellobiohydrolase," *Sci. China C Life Sci*. 47(1), 18-24. https://doi.org/10.1360/02yc0163

Zhao, S., Yan, Y. S., He, Q. P., Yang, L., Yin, X., Li, C. X., Mao, L. C., Liao, L. S., Huang, J. Q., Xie, S. B., *et al*. (2016). "Comparative genomic, transcriptomic and secretomic profiling of *Penicillium oxalicum* HP7-1 and its cellulase and xylanase hyper-producing mutant EU2106, and identification of two novel regulatory genes of cellulase and xylanase gene expression," *Biotechnol Biofuels* 9, article 203. https://doi.org/10.1186/s13068-016-0616-9

Zheng, F., Basit, A., Wang, J., Zhuang, H., Chen, J., and Zhang, J. (2024). "Characterization of a novel acidophilic, ethanol tolerant and halophilic GH12 β-1,4-endoglucanase from *Trichoderma asperellum* ND-1 and its synergistic hydrolysis of lignocellulosic biomass," *Int. J. Biol. Macromol.* 254(Pt 1), 127650. https://doi.org/10.1016/j.ijbiomac.2023.127650
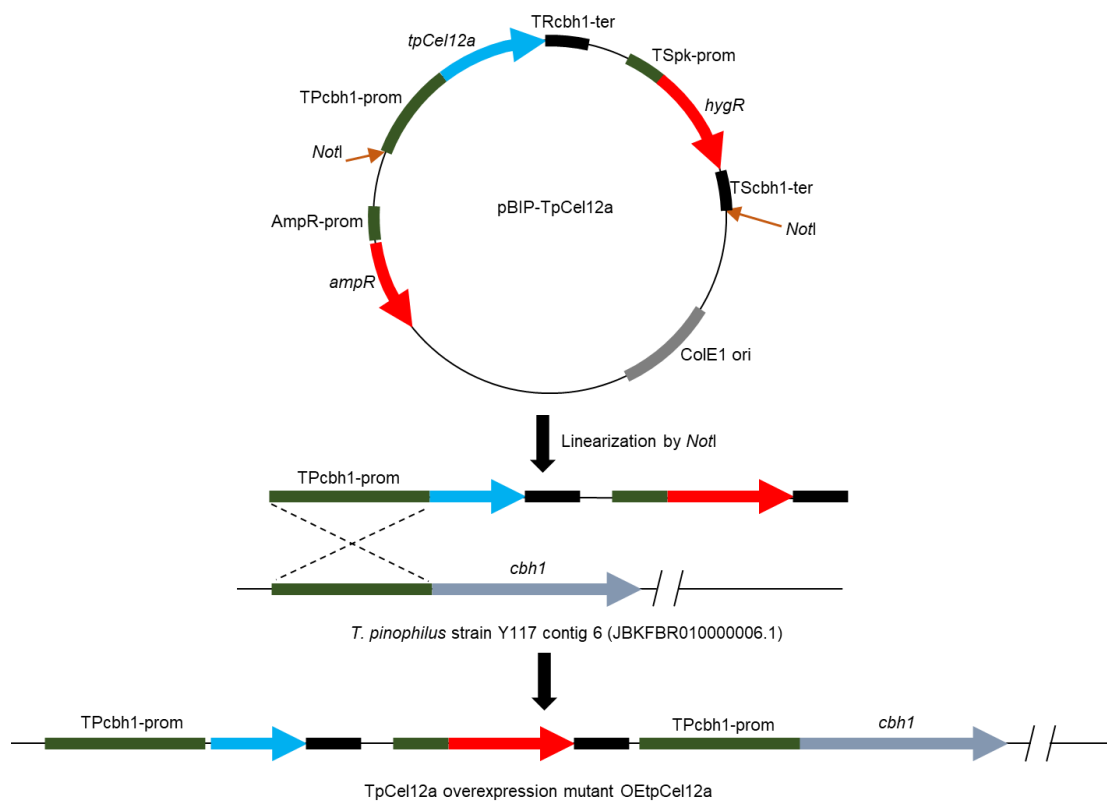
APPENDIX

Appendix I. Strains and Plasmids Used in this Study

Strains/Plasmids	Features	Origin
E. coli TOP10	F⁻*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Thermo Fisher Scientific
T. pinophilus strain Y117	Cellulase high-yielding strain	Lab
OEtpCel12a	TpCel12a overexpression mutant	This study
OEtpCel12a(+TQA)	TpCel12a(+TQA) overexpression mutant	This study
pUC57	*ColE1* origin, Amp^R	Lab
pLKO.1-Hygro	f1 origin, Amp^R, Hyg^R	Lab
pBIP-TpCel12a	Overexpressing recombinant TpCel12a	This study
pBIP-TpCel12a(+TQA)	Overexpressing TpCel12a mutant TpCel12a(+TQA)	This study

Appendix II. Oligonucleotides Used in this Study

primer	Sequence (5'→3')
TPcbh1-prom-f	aaacgacggccagtgaattc**GCGGCCGC**GGTCGAATAGCTGCGACATATA
TPcbh1-prom-r	aaaagttagcttcatTGTGTCGATTGCTTCTGACTGT
tpcel12a-f	gaagcaatcgacacaATGAAGCTAACTTTTCTCCTGA
tpcel12a-r	attgcgtatatggaaTTAATGGTGGTGATGGTGATGGTGGTGATTGACAGATGCAGACCAATG
TRcbh1-ter-f	catcaccaccattaaTTCCATATACGCAATAGTATTC
TRcbh1-ter-r	gcgaggtaccgagctcTCGAACAAAGCAAGTCGGGCGA
TRpk-prom-f	tagatatcggatcccgggAGCCTGCTGGCTGTTTCGCCAG
TRpk-prom-r	ttcaggcttcttcatCTTGGAAGTTGCTGCGGGGAAG
hyg-f	gcagcaacttccaagATGAAGAAGCCTGAACTCACCG
hyg-r	accgatatcaaccatCTATTCCTTTGCCCTCGGACGA
TScbh1-ter-f	agggcaaaggaatagATGGTTGATATCGGTTGGATGG
TScbh1-ter-r	atgcaggcctctgcag**GCGGCCGC**GGTAATCAATTGTCTTTTTTTTTAGG
TpCel12a-inTQA-f	GTAAACACCCAAGCCGGATCCCAAAAAACGTACAG
TpCel12a-inTQA-r	GGATCCGGCTTGGGTGTTTACGCCATACCACAG
Bold and underlined nucleotides indicate the recognition site of *Not*I and codons of 8×His-tag, respectively	

Appendix III: Methods

Sample preparation for liquid chromatography-tandem mass spectrometry (LC-MS/MS)

The purified TpCel12a was treated with 10 mmol/L dithiothreitol (DTT) at 56 °C for 1 h to reduce its disulfide bonds. Subsequently, the reduced thiol groups were alkylated by adding iodoacetamide (IAA) at a final concentration of 20 mmol/L in the dark at room temperature. The alkylated TpCel12a was then subjected to enzymatic digestion using Glu-C protease under sealed conditions at 37 °C for 16 h. After digestion, a peptide extraction solution (5% trifluoroacetic acid [TFA], 50% acetonitrile [ACN], and 45% water) was added, followed by incubation in a 37 °C water bath for 1 h. The mixture was then ultrasonicated for 5 min and centrifuged for 5 min to obtain the peptide extract. Finally, the digested peptides of TpCel12a were desalted and prepared for further analysis.

LC-MS/MS analysis of digests

LC-MS/MS was conducted using an Orbitrap Eclipse mass spectrometer (Thermo Scientific, USA) coupled with an Ultimate 3000 nano-LC system. For each injection, 5 μL of the sample was loaded onto a C18 PepMap100 trap column (300 μm × 5 mm, Thermo Scientific, USA) and subsequently separated on an analytical column (Acclaim PepMap RPLC, 150 μm × 15 cm, Thermo Scientific, USA) using a 72-min linear gradient. The gradient profile was as follows: 4%–9% solvent B (0.1% formic acid in 90% acetonitrile) over 3 min, 9%–22% B over 22 min, 22%–32% B over 22 min, 32%–42% B over 13 min, 42%–95% B over 4 min, followed by a 8-min hold at 95% B (solvent A: 0.1% formic acid in water). The flow rate was maintained at 0.3 μL/min. Mass spectrometry analysis was performed in data-dependent acquisition (DDA) mode. Full-scan MS spectra (350 to 1500 m/z) were acquired at a resolution of 60,000 (at 200 m/z), with an automatic gain control (AGC) target of $4×105$ ions and a maximum injection time of 50 ms. The top 15 or 20 most intense precursor ions were selected for fragmentation via higher-energy collisional dissociation (HCD) at a normalized collision energy (NCE) of 30%. MS/MS spectra were collected at a resolution of 15,000, with an AGC target of $5×104$ ions and a maximum injection time of 22 ms. The peptide identification was finally achieved by searching the MS/MS spectra against the target sequence.

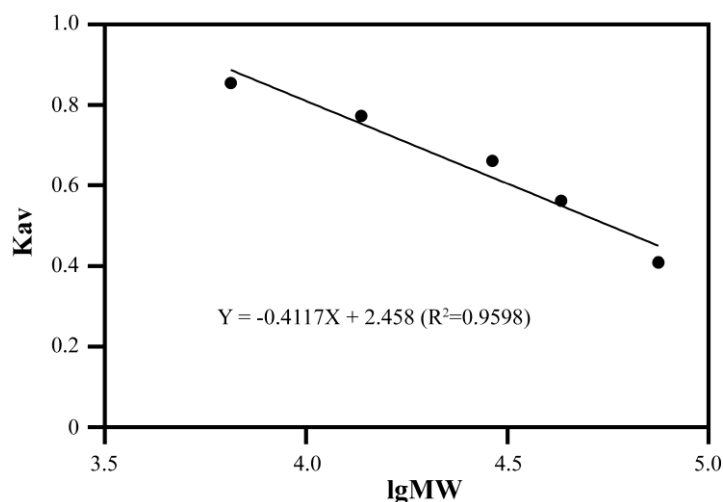*MALDI-TOF MS analysis of final products of xyloglucan catalyzed by TpCel12a*

MALDI-TOF MS was conducted on a Bruker Autoflex speed MALDI-TOF MS spectrometer equipped with a 337 nm nitrogen laser. Generally, the enzymatic hydrolysates were dissolved in ultrapure water. A saturated matrix solution was prepared by dissolving 2,5-dihydroxybenzoic acid (DHB) in 50% (v/v) methanol aqueous solution containing 0.1% (v/v) trifluoroacetic acid (TFA). For MALDI target preparation, 1 μL of the aqueous hydrolysate solution was first spotted onto a polished stainless steel target plate and allowed to air-dry at room temperature. Subsequently, 1 μL of the DHB matrix solution was carefully overlaid onto the dried sample spot and similarly air-dried. Mass spectrometric analysis was performed in positive ion reflection mode. The instrument parameters were optimized for optimal signal-to-noise ratio in the target mass range. Raw spectral data were processed and analyzed using flexAnalysis software (Bruker Daltonics) with appropriate baseline subtraction and noise reduction algorithms applied.

Appendix IV. Construction of TpCel12a Overexpression Mutant OEtpCel12a

Plasmid pBIP-TpCel12a was linearization by *Not*I. Then the linear fragment was introduced into *T. pinophilus* Y117 by protoplast transformation. Hygromycin resistance clones were selected and the expression of TpCel12a was confirmed by western blot.

Appendix V. Calibration Curve of Kav for the Standard Proteins (Low Molecular Kit)

$Y = -0.4117X + 2.458 \ (R^2=0.9598)$

Appendix VI. MALDI-TOF MS Analysis of Hydrolysis Products of CMC
Degraded by TpCel12a