Cloning and characterization of thermophilic cellulase and its application in the transformation of ginsenosides

Fei Zheng  
Changchun University of Chinese Medicine

Huanxi Zhao  
Changchun University of Chinese Medicine

Nan Wang  
Changchun University of Chinese Medicine

Peng Zhong  
Changchun University of Chinese Medicine

Kailu Zhou  
Changchun University of Chinese Medicine

Shanshan Yu (✉ yushanshan001@aliyun.com)  
Changchun University of Chinese Medicine  https://orcid.org/0000-0001-6532-7205

Research Article

Keywords: Biotransformation, Cellulase, Ginsenoside, Glycoside hydrolase family, Fervidobacterium pennivorans DSM9078

Posted Date: June 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1602183/v1

License: ☺️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

A novel cellulase (BcelFp) was identified from *Fervidobacterium pennivorans* DSM9078 which had biotransformation activity for PPD-type ginsenosides. Sequence analysis of BcelFp revealed that it could be classified into glycoside hydrolase family 5 (GH5). The gene encoding a 323-amino acid protein was cloned and expressed in *Escherichia coli*. The recombinant enzyme was purified, and its molecular weight was approximately 37 kDa. The recombinant BcelFp exhibited an optimal activity at 95 °C and pH 5.5 and showed high thermostability. The cellulase had high selectivity for cleaving the outer glucose moiety at the C3 carbon of ginsenoside Rb1, Rb2, Rc and Rd, which produced the more pharmacologically active gypenoside XVII (GypXVII), Compound O (CO), Compound Mc1 (CMc1) and F2, respectively. The $K_m$ values for Rb1, Rb2, Rc and Rd were 3.66 ± 0.04 µM, 4.02 ± 0.12 µM, 5.95 ± 0.03 µM, 0.67 ± 0.006 µM, respectively. The $k_{cat}/K_m$ value of BcelFp for ginsenoside Rd was 27.91 mM$^{-1}$s$^{-1}$, which was much higher than the previously enzymes. This study was the first report of the highly efficient and selective transformation of GypXVII, CO, CMc1 and F2 from Rb1, Rb2, Rc and Rd by a GH5-family thermophilic cellulase.

Introduction

Ginseng, the root of the *Panax* ginseng C. A. Mey., has been widely used as a kind of valuable Chinese-traditional medicine in the Eastern Asian for thousands of years. Ginsenosides are the principal components responsible for the diverse and significant effects of ginseng. The pharmacological activities include anti-tumor (Xu et al. 2021), anti-inflammatory (Yi 2021), anti-diabetic (Zhou et al. 2019), anti-fatigue (Arring et al. 2018), anti-aging (Shen et al. 2017) and anti-microbial activities (Wang et al. 2020).

More than 180 ginsenosides presenting different pharmaceutical activities have been identified from Ginseng. According to the type of the aglycone structure, they are divided into two groups: protopanaxadiol (PPD)-type ginsenosides and protopanaxatriol (PPT)-type ginsenosides. These ginsenosides comprise a nonsugar constituent of dammarane skeleton and a sugar constituent including 1–4 molecule glycosides such as glucose, arabinopyranose, arabinofuranose, xylose, and rhamnose. Structural differences such as linked positions, inner and outer residues and types of sugar moieties in ginsenosides contribute to their different pharmacological activities. Deglycosylated ginsenosides are proven to be more pharmaceutically active because of their smaller size, better permeability across the cell membrane and higher bioavailability (Ku 2016). In our previous work, we demonstrated that deglycosylated ginsenosides (Rd, GypXVII, and PPT) had significantly greater anti-inflammatory activity than their glycosylated precursors (Rb1, Re and Rg1) (Yu et al. 2017). As deglycosylated ginsenosides are low in content or even absent in ginseng, transformation of deglycosylated ginsenosides through hydrolysis of the sugar moieties in glycosylated ginsenosides have attracted wide attention recently.

Various transformation approaches such as physiochemical methods including heating, acid treatment and alkali treatment, and bioconversion methods using microorganisms and enzymes have been
attempted (Park et al. 2010). Among these methods, the enzymatic conversion is the most promising method because of its high substrate specificity and stability, low levels of by-products, and high production yields. To minimize processing time and production cost, recombinant enzymes obtained from E. coli strains are commonly applied for the ginsenoside conversion. For instance, the recombinant β-glucosidase from Bifidobacterium breve ATCC 15700 was identified to produce CK from ginsenoside Rd (Zhang et al. 2019). Higher productivity was achieved when recombinant thermophilic enzymes were used to transform ginsenosides. The β-glycosidase from Sulfolobus acidocaldarius could convert ginsenoside Rb1, Rb2, Rc, and Rd with a mole yield of 99% after 24 h (Noh et al. 2009).

Until now, many glycosidases have been explored to hydrolyze the sugar moieties linked to the C3, C6, and C20 positions in glycosylated ginsenosides. These ginsenoside-transforming glycosidases including β-glucosidases, arabinopyranosidases, arabinofuranosidases, xylosidases, rhamnosidases and β-galactosidases belong to GH family 1, 2, 3, 39, 42, 51, and 78 (Shin and Oh 2015). For instance, The GH1 β-glucosidase from Sphingopyxis alaskensis specifically hydrolyzed the outer glucose at the C3 position in PPD-type ginsenosides (Zhang et al. 2019). Thermophilic glycosidases from GH family 1, 3, 39 and 42 were also identified to transform PPD- or PPT-type ginsenosides (Noh et al. 2009; Pei et al. 2015; Xie et al. 2015; Shin et al. 2013). Considering the complex and diverse structures of ginsenosides, the number of ginsenoside-transforming glycosidases is still limited. And the biotransformation activity, specificity, and thermostability of most enzymes still do not meet the industrial demands. It is therefore meaningful to explore novel glycosidases with good thermostability, high catalytic efficiency and specificity.

The concentration of GypXVII, CO, CMc1, and F2 are quite low in Ginseng. Whereas these four deglycosylated ginsenosides exhibit various pharmaceutical activities. GypXVII exerts strong cardio-protective (Yu et al. 2019), anti-apoptotic and autophagic activities (Sun et al. 2021) and protects against reperfusion and spinal cord injury (Luo et al. 2021). Ginsenoside F2 possesses anti-obesity (Siraj et al. 2015) and anti-cancer activity (Shin et al. 2012). There are very few researches about pharmaceutical activities of CO and CMc1 due to the production difficulty. GypXVII, CO, CMc1, and F2 can be generated by hydrolysis of the outer glucose linked to the C3 position of ginsenoside Rb1, Rb2, Rc, and Rd, respectively. Therefore, the aim of our work is to transform glycosylated ginsenosides to the more pharmaceutically active ginsenosides by enzymatic catalysis with high efficiency and productivity.

In this paper, a novel thermophilic cellulase from F. pennivorans DSM9078 was cloned, purified, and characterized. It showed high specificity that hydrolyzed only the outer glucose at the C3 position in PPD-type ginsenosides, and biotransformed the glycosylated ginsenoside Rb1, Rb2, Rc, and Rd to the deglycosylated ginsenoside GypXVII, CO, CMc1, and F2, respectively. This study was the first report of the highly efficient and selective transformation of PPD-type ginsenosides by a GH5-family thermophilic cellulase.

Materials And Methods

Materials
*Escherichia coli* strains were incubated at 37 °C in Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) with 50 mg/L kanamycin. Authentic ginsenoside Rb1, Rb2, Rc, Rd, GypXVII and F2 were purchased from Shanghai Yuanye Biological Technology Co. Ltd. (Shanghai, China).

**Phylogenetic analysis of BcelFp**

The theoretical molecular weight (Mw) and isoelectric point (pI) of BcelFp were estimated on the ExPasy server (http://web.expasy.org/compute pi/) Gasteiger et al. 2005. Homologs of BcelFp (Genbank No. AFG35892.1) were searched with the BLASTp program on NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The glycoside hydrolases in GH5 family were chosen at the CAZy web (http://www.cazy.org/fam/GH5.html), and the sequence information of those proteins were collected by means of the CAZy web page links. Multiple alignments of BcelFp and three characterized glycosidases from GH5 family were performed using the ClustalX program (Larkin et al. 2007). The phylogenetic tree of BcelFp and the enzymes chosen from GH5 family was constructed using the neighbor-joining method (Saitou and Nei 1987) with default parameters and 1000 bootstrap in the MEGA4 Program (Felsenstein 1985).

**Molecular cloning, expression and purification of BcelFp**

Genomic DNA from *F. pennivorans* DSM9078 was extracted using a genomic DNA extraction kit (TIANGEN, Beijing, China) and used as a template of the gene cloning. The cellulase gene, *bcelfp*, from *F. pennivorans* DSM9078 was amplified via polymerase chain reaction (PCR). The primers were designed based on genomic sequence (Genbank No. WP_245530410.1): upstream (5’-CAGCAGGGATCCATGGATCAGTCAGTTGCT-3’) and downstream (5’-CAGCAGCTCGAGTTATTCTTTGCTTTCTCACA-3’) with BamHI and XhoI restriction sites (underlined), respectively. The PCR amplified DNA fragment with a His-tag at N terminal was purified and inserted into the pET28a vector digested with BamHI and XhoI. The recombinant pET28a-*bcelfp* was transformed into *E.coli* BL21 (DE3), which was induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) at 16 °C for an additional 12 h.

The recombinant strains were collected by centrifugation at 6,000×g for 20 min and resuspended in a lysis buffer (50 mM Tris-HCl, pH 7.1). The cells were then sonicated and centrifuged at 14,000×g for 30 min at 4 °C to remove the debris. The supernatants containing the target proteins were loaded onto a NiNTA affinity chromatography column (GE Healthcare) and purified using a 20-100 mM imidazole gradient. The purified enzyme was dialyzed against 50 mM phosphate-citrate buffer (pH 6.0) and concentrated to 1.0 mg/mL. The protein homogeneity was confirmed by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

**Cellulase activity assay of BcelFp**
The hydrolytic activity of the cellulase BcelFp was assayed according to the Ghose (Ghose 1987) and DNS (Miller 1959) methods. Protein concentrations were determined by using the Bradford Protein Assay Kit (Sangon Biotech, Shanghai, China). The reaction mixture composed of 1% CMC (w/v) and 1 mg/ml cellulase BcelFp in 50 mM phosphate-citrate buffer (pH 6.0). After incubation at 95 °C for 5 min, DNS was added to terminate the action, and the mixture was boiled in 100 °C water for 5 min. The absorption of the reaction mixture was measured at 540 nm. One unit (IU) of enzyme activity was defined as the amount of enzyme that released 1 μmol reducing sugars per min.

**Enzyme characterization of BcelFp**

Using CMC as the substrate, the effects of temperature and pH on enzyme activity were investigated by assaying the cellulase activity according to the method described previously. The pH optima of BcelFp was tested over the pH range of 4.0-8.0 in 50 mM phosphate-citrate buffer. The temperature optima of BcelFp was measured between 30-100 °C (5 °C intervals) at the optimum pH. Thermal stability of BcelFp was studied by incubating about 1.5 mg/mL purified enzyme solutions at 85 °C, 90 °C or 95 °C for various lengths of time in 50 mM phosphate-citrate buffer (pH 6.0). The residual activity on CMC was determined at 95 °C in 50 mM phosphate-citrate buffer (pH 6.0). pH stability of BcelFp was tested by incubating about 1.5 mg/mL purified enzyme solutions in pH 4, pH 5, pH 6 and pH 7 for various lengths of time at 85 °C. The residual activity on CMC was determined at 95 °C in 50 mM phosphate-citrate buffer (pH 6.0).

The effects of metal ions (NH$^{4+}$, Na$^{+}$, K$^{+}$, Co$^{2+}$, Ca$^{2+}$, Zn$^{2+}$, Ba$^{2+}$, Mn$^{2+}$, and Fe$^{3+}$), and EDTA on enzyme activity were investigated. The enzyme activity with no addition for the control was set as 100%. After incubation the enzyme with various metal ions (5 mM and 10 mM), and EDTA (5 mM and 10 mM) at room temperature for 30 min, the residual activity was measured according to the method described previously using CMC as substrate at pH 6.0 and 95 °C.

**Differential scanning calorimetry (DSC)**

Thermal inactivation of BcelFp was analyzed by DSC on a VP-Capillary differential scanning calorimeter (MicroCal, LLC, GE Healthcare) over a temperature range from 35 to 120 °C. The purified enzyme was dialyzed against 50 mM PBS buffer (pH 6.0) and concentrated to 1.5 mg/mL. The corresponding buffer was used as a reference. The equilibrated enzyme was scanned at a rate of 2.0 K/min. The scans were analyzed after subtraction of an instrument baseline recorded with buffer in both cells using the software package Origin provided by the manufacturer.

**Substrate specificity of BcelFp**

The substrate specificity of BcelFp was investigated by using the following polysaccharide substrates: CMC (carboxymethyl cellulose sodium salt, medium viscosity; Fluka), RAC (regenerated amorphous cellulose) (Zhang et al. 2006), Avicel (PH-101; Fluka), β-glucan from barley.
(Sigma), laminarin (Sigma), soluble starch (Sigma), and pustulan (Sigma). The enzyme activities using these polysaccharide substrates were measured according to the method described previously. One unit (IU) of enzyme activity was defined as the amount of enzyme that released 1 μmol reducing sugars per min.

The p-nitrophenyl β-d-glucopyranoside (pNPG) was used to test the β-glucosidase activity of BcelFp. Hydrolysis of pNPG was measured at 95 °C in 50 mM phosphate–citrate buffer (pH 6.0). The activity was determined by measuring the increase in absorbance at 405 nm due to the release of pNP. One unit (IU) of activity was defined as the amount of enzyme liberating 1 μmol of p-nitrophenol per min.

The substrate specificity of BcelFp towards different PPD- and PPT-type ginsenosides were also studied. Enzyme solution (2 mg/mL) was reacted with equal volume of ginsenoside solution (1 mg/mL) in 50 mM phosphate–citrate buffer (pH 6.0) at 95 °C for 5-30 min. The ginsenoside-tranforming activity of BcelFp on different ginsenosides were analyzed by HPLC. One unit (IU) of activity was defined as the amount of enzyme catalyzing the conversion of 1 μmol of ginsenoside substrate per min.

**Biotransformation of PPD-type ginsenosides by BcelFp**

The biotransformation ability of recombinant BcelFp towards PPD-type ginsenosides was studied. Ginsenoside Rb1, Rb2, Rc, and Rd were used as substrates. Enzyme solution (2 mg/mL) was reacted with equal volume of ginsenoside solution (1 mg/mL) in 50 mM phosphate–citrate buffer (pH 6.0) at 95 °C. The reaction solution without enzyme served as blank control. Samples were obtained at regular intervals and analyzed via HPLC.

Kinetic studies were performed by using Rb1, Rb2, Rc and Rd at concentrations from 0.5 to 10.0 mM in 50 mM phosphate–citrate buffer (pH 6.0) at 95 °C. One unit (IU) of activity was defined as the amount of protein required to convert 1 μmol of substrate per min. All assays were performed in triplicate. The parameters K_m and V_max were determined using the enzyme kinetics program described by Cleland (Cleland 1979).

**HPLC-Q-TOF-MS/MS analysis**

Chromatographic separation was performed on an Agilent 1200 series HPLC system (Agilent Technologies, Santa Clara, CA, USA) coupled with a Syncronis C18 chromatographic column (5 cm×3.0 mm, 2.7 μm, Supelco, USA). The column oven temperature was maintained at 35 °C, and the mobile phases A and B were water with 0.1% formic acid and acetonitrile, respectively. The gradient elution program was determined as follows: 0-2.5 min, 19% (B); 2.5-5 min, 19-30% (B); 5-11 min, 30-33% (B); 11-20 min, 33-45% (B); 20-25 min, 45-65% (B). The injection volume was 5 μL, and the flow rate was 0.4 mL/min.

Mass spectrometric detection was carried out on an Agilent 6520 Q-TOF-MS/MS (Agilent Technologies, Santa Clara, CA, USA) equipped with electrospray ionization source operated. The scan range for MS
acquisition was from 100 to 2200 \textit{m/z} in negative ionization mode. The parameters of ion source were set as follows: drying gas, N$_2$; flow rate, 4.0 mL/min; drying gas temperature, 350 $^\circ$C; nebulizer, 30 psig; capillary voltage, 3500 V; fragmentor, 350 V; cone voltages, 65 V. Data analysis was performed by Mass Hunter Qualitative (MHQ) software, version B.03.01 (Agilent Technologies, Santa Clara, CA, USA).

$^{13}$C NMR analysis

The biotransformed products were isolated using a Waters prep-HPLC system (Waters, USA) with a Waters Sunfire Prep C$_{18}$ OBD$^\text{TM}$ column (19 mm×50 mm, 5 µm). Isocratic elution was performed at room temperature using a mixture of H$_2$O/ACN (56:44, v/v) for product 1, H$_2$O/ACN (48:52, v/v) for product 2, H$_2$O/ACN (54:46, v/v) for product 3 and H$_2$O/ACN (54:46, v/v) for product 4, respectively. The flow rate was 15 mL/min, and the UV detection was at 203 nm. The injection volume was set at 200 µL. The purity of the products were evaluated by HPLC in an Agilent 1200 series HPLC system (Agilent Technologies, USA). The separation conditions were the same as that used in the HPLC analysis. The $^{13}$C NMR spectra of the purified products were recorded on a Bruker AVIII 600 MHz spectrometer at an operating frequency of 151 MHz (Bruker BioSpin, Germany).

Results

Sequence analysis of BcelFp from \textit{F. pennivorans} DSM9078

The cellulase gene \textit{BcelFp} consisted of 969 bp encoding 323 amino acids with a theoretical molecular mass of 37.89 kDa and a theoretical pI value of 5.43. The amino acid sequence of BcelFp (Genbank No. AFG35892.1) exhibited highest similarity with GH 5 proteins from \textit{Fervidobacterium islandicum} (85.7% identity, Genbank No. WP_052107242.1 ) and \textit{Fervidobacterium changbaicum} (85.7% identity, Genbank No. WP_090223359.1). These proteins have not yet been characterized. The nearest characterized glycoside hydrolase (76% identity, GenBank No. WP_011994708.1) in the CAZy database was cellulase FnCel5A from \textit{Fervidobacterium nodosum}. The BcelFp and FnCel5A were both from thermophilic bacteriums belonging to genus \textit{Fervidobacterium}. FnCel5A was the first cellulase of the genus \textit{Fervidobacterium} that has been cloned and expressed (Wang et al. 2010). The alignment of BcelFp with several characterized glycoside hydrolases from GH5 indicated that these proteins shared some conserved peptide motifs, namely NEP (residues 143-145), HYH (residues 203-205) and GEFH (residues 259-262) (Fig. 1). The Glu144 and Glu260 residues were typical catalytic residues of the GH5 enzymes, which confirmed that BcelFp belonged to GH5 family.

Glycosyl hydrolases are classified into different families according to their amino acid sequence similarities (http://www.cazy.org/fam/acc_GH.html). Glycoside hydrolase family 5, one of the largest GH families, was subdivided into 53 subfamilies (Aspeborg et al. 2012). In order to gain a better understanding of the evolutionary position of BcelFp in glycoside hydrolase family 5, we constructed the phylogenetic tree using the neighbor-joining method in the MEGA4 program with bootstrap values based on 1,000 replications. The resulting consensus tree is presented in Fig. 2. BcelFp from \textit{F. pennivorans}
DSM9078 clustered within subfamily 25 and formed a separate, well-supported clade with cellulase from *A. thermocellus* and endoglucanase (FnCel5A) from *F. nodosum* Rt17-B1.

**Expression, purification and characterization of the cellulase BcelFp**

The putative cellulase gene from *F. pennivorans* DSM9078 was cloned and expressed in *Escherichia coli* under the control of the IPTG-inducible promoter T7. After being induced under 16 °C for 12 h with 1 mM IPTG, the recombinant enzyme was solubly overexpressed in *E.coli* cells. The recombinant BcelFp was purified by His-trap affinity chromatography with a final purification of 2.3-fold and a specific activity of 297 U/mg for CMC. The expressed enzyme was determined as a single band by SDS-PAGE, with a molecular mass of approximately 38 kDa (See Additional file 1: Fig. S1), which was almost consistent with the molecular weight of 37,892 Da calculated from 323 amino acids.

The optimum pH and temperature of the purified BcelFp were determined using CMC as the substrate. The maximum activity was observed at pH 6.0, and the cellulase activity was higher than 50% of the maximum activity at the pH range from 4.0 to 6.5. While from pH 6.5, the enzyme activity decreased swiftly, representing that the enzyme was active over narrow acid pH range (Fig. 3A). After being incubated at various pHs (pH 4, pH 5, pH 6 and pH 7) for 1 h, more than 90% of enzyme activity remained at acid pHs, while 70% of enzyme activity remained at pH 7.0, indicating that BcelFp was more stable at acid pHs (Fig. 3B). The optimal temperature for BcelFp activity was 95 °C, while it also displayed high activity between 60 and 100 °C (Fig. 3C). Hence, the BcelFp was thermophilic enzyme.

The effect of temperature on enzyme stability was investigated using CMC as the substrate. As seen in Fig. 3D, the enzyme was incubated for various lengths of time at 85, 90 and 95 °C at pH 6.0 and then the residual activities were measured. The enzyme was very stable at high temperature, and its half-lives at 85, 90 and 95 °C were 60, 25, and 10 h, respectively. When DSC analysis was performed to determine the thermal transition mid-point (Tm) of BcelFp, it was 96 °C (Fig. 3E).

The effects of metal ions, and EDTA on BcelFp activity were also investigated (Table 1). Na+ and Mn2+ were able to significantly increase enzyme activity under the high concentration of 10 mM. Moreover, Fe3+, Co2+, Zn2+ were able to inhibit enzyme activity under concentrations of 5 mM and 10 mM. No significant effect was observed in the presence of Ca2+, NH4+, Ba2+, and K+ under the high concentration of 10 mM. The chelating agent EDTA did not inhibit the activity, which indicated that BcelFp was not metalloprotein.

**Table 1** Effects of metal ions and EDTA on the enzyme activity of BcelFp
<table>
<thead>
<tr>
<th>Additives</th>
<th>Relative Activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>100</td>
</tr>
<tr>
<td>Metal ions</td>
<td></td>
</tr>
<tr>
<td>NH₄Cl</td>
<td>102.81</td>
</tr>
<tr>
<td>NaCl</td>
<td>108</td>
</tr>
<tr>
<td>BaCl₂</td>
<td>78.28</td>
</tr>
<tr>
<td>KCl</td>
<td>92.34</td>
</tr>
<tr>
<td>MnCl₂</td>
<td>95.66</td>
</tr>
<tr>
<td>CaCl₂</td>
<td>103</td>
</tr>
<tr>
<td>ZnCl₂</td>
<td>90.96</td>
</tr>
<tr>
<td>CoCl₂</td>
<td>85.06</td>
</tr>
<tr>
<td>FeCl₃</td>
<td>32.16</td>
</tr>
<tr>
<td>Inhibitors</td>
<td></td>
</tr>
<tr>
<td>EDTA</td>
<td>87.82</td>
</tr>
</tbody>
</table>

**Substrate specificity of the cellulase BcelFp**

The substrate specificity of BcelFp was investigated using substrates including CMC, barley β-glucan, RAC, Avicel, Laminarin, soluble starch, pustulan, and pNPG. CMC is typical substrate for determining endoglucanase activity, while Avicel is used for measuring exoglucanase activity. The pNPG is a synthetic compound which serves as an optimal substrate for β-glucosidase. As seen in Table 2, BcelFp displayed high activity towards CMC, but undetectable activity on Avicel and pNPG, therefore it was an endoglucanase cellulase. Moreover, BcelFp was able to hydrolyze RAC, structure of which was similar to CMC. In addition, selectivity of BcelFp on β-1,6-, β-1,3-, β-1,4-, β-1,6- linkages in glycosides were also tested. Barley β-glucan is composed of a mixed-linked β-1,3/1,4-glucans. Laminarin contains primarily β-1,3 and a portion of β-1,6-glycosidic bonds. Soluble starch consists of β-1,4/1,6-glucans. Pustulan is a linear β-1,6-linked glucans. BcelFp exhibited highest activity on barley β-glucan, while no activity on laminarin, soluble starch, and pustulan, which further indicating that BcelFp can randomly hydrolyze the β-1,4-glucopyranosyl linkage, but cannot act on β-1,6-, β-1,3-, β-1,4-, β-1,6- linkages in glycosides.

**Table 2** Enzyme specificity for BcelFp on various substrates
The substrate specificity of BcelFp on PPD- and PPT-type ginsenosides was also measured. As seen in Table 3, the sugar moieties including glucose, arabinopyranose, and arabinofuranose linked to C3 and C20 position in PPD-type ginsenosides. And the sugar moieties including glucose, xylose, and rhamnose linked to C6 and C20 position of PPT-type ginsenosides. The relative activity of BcelFp for the PPD-type ginsenosides was in the order Rd>Rb1>Rb2>Rc. However, BcelFp showed undetectable activity on GypXVII, CO, CMc1, F2, and Rg3. In addition, for PPT-type ginsenosides as substrates, the enzyme exhibited no activity on R1, Re, Rg1, Rg2 and Rh1, indicating that BcelFp didn't hydrolyze sugar moieties linked to the C6 and C20 position of PPT-type ginsenosides.

**Table 3** Substrate specificity of BcelFp on PPD- and PPT-type ginsenosides

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Relative Activity(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMC</td>
<td>100&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Barley β-glucan</td>
<td>486</td>
</tr>
<tr>
<td>RAC</td>
<td>21</td>
</tr>
<tr>
<td>Avicel</td>
<td>ND&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Laminarin</td>
<td>ND</td>
</tr>
<tr>
<td>Soluble starch</td>
<td>ND</td>
</tr>
<tr>
<td>pustulan</td>
<td>ND</td>
</tr>
<tr>
<td>pNPG</td>
<td>ND</td>
</tr>
</tbody>
</table>

<sup>a</sup> The activity against CMC was assumed to be 100%, and corresponded to a specific activity of 297 U/mg.

<sup>b</sup> Not detected, specific activity is not detected by the analytical methods used in this study.
<table>
<thead>
<tr>
<th>Type</th>
<th>Substrate&lt;sup&gt;a&lt;/sup&gt;</th>
<th>C3</th>
<th>C6</th>
<th>C20</th>
<th>Relative Activity (%)&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPD-</td>
<td>Rb1</td>
<td>Glu(1→2)Glu-</td>
<td>Glu(1→6)Glu-</td>
<td>68.18±1.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rb2</td>
<td>Glu(1→2)Glu-</td>
<td>Arap(1→6)Glu-</td>
<td>27.27±3.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rc</td>
<td>Glu(1→2)Glu-</td>
<td>Araf(1→6)Glu-</td>
<td>25.18±0.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rd</td>
<td>Glu(1→2)Glu-</td>
<td>Glu-</td>
<td>100±1.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rg3</td>
<td>Glu(1→2)Glu-</td>
<td>H-</td>
<td>ND&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GypXVII</td>
<td>Glu-</td>
<td>Glu(1→6)Glu-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C0</td>
<td>Glu-</td>
<td>Arap(1→6)Glu-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CMc1</td>
<td>Glu-</td>
<td>Araf(1→6)Glu-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>Glu-</td>
<td>Glu-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>PPT-</td>
<td>R1</td>
<td>Xyl(1→2)Glu-</td>
<td>Glu-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Re</td>
<td>Rha(1→2)Glu-</td>
<td>Glu-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rg1</td>
<td>Glu-</td>
<td>Glu-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rg2</td>
<td>Rha(1→2)Glu-</td>
<td>H-</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rh1</td>
<td>Glu-</td>
<td>H-</td>
<td>ND</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Final concentration of substrate was 1.0 mM.

<sup>b</sup> The activity against Rd was assumed to be 100%, and corresponded to a specific activity of 10.8 U/mg.

<sup>c</sup> Not detected, specific activity is not detected by the analytical methods used in this study.

**Biotransformation of PPD-type ginsenosides by the cellulase BcelFp**

For the verification of the biotransformation pathways of the four PPD-type ginsenosides (Rb1, Rb2, Rc, and Rd) by using BcelFp, the HPLC analyses were performed. As shown in Fig. 4, ginsenoside Rb1 Rb2, Rc, and Rd had decomposed dramatically after 1 h of reaction. With longer reaction time, still only one product was detected for each substrate (data not shown). Peak 1 and 4 were identified to be GypXVII and F2 by comparison of their retention time with standards, respectively. As lack of standards for peak 2 and 3, their structures were further analyzed by HPLC-MS. As seen in Additional file 1: Fig. S2, The molecular weight of peak 2 was calculated to be 916 u based on its [M-H]<sup>-</sup> ion at m/z 915. The difference between 1078 u (Rb2) and 916 u (peak 2) is 162 u, which corresponds to a glucose moiety. So, BcelFp catalyzed the release of one glucose residue from Rb2 to generate product 2. As Rb2 has an outer glucose linked to C3 position and an outer arabinopyranose linked to C20 position, the deglycosylation
should occur at C3 position of Rb2 to generate CO. Similarly, peak 3 was identified to be CMc1 (See Additional file 1: Fig. S3). Combination with the $^{13}$C NMR analysis, product 1-4 were finally confirmed to be GypXVII, CO, CMc1 and F2 (See Additional file 1: Table. S1), respectively. Based on the structural analysis, GypXVII, CO, CMc1, and F2 were generated by hydrolysis of the outer glucose linked to the C3 position of ginsenoside Rb1, Rb2, Rc and Rd, respectively. The biotransformation pathways of ginsenoside Rb1, Rb2, Rc and Rd catalyzed by BcelFp were presented in Fig. 5. In addition, GypXVII, CO, CMc1, and F2 were not further hydrolyzed. Hence, BcelFp didn’t hydrolyze sugar moieties linked to C20 position and the inner glucose residue at C3 position of PPD-type ginsenosides.

As seen in Table 4, the Michaelis-Menten constants ($K_m$), Maximum Velocity ($V_{max}$), and catalytic efficiencies ($k_{cat}/K_m$) for Rb1, Rb2, Rc, and Rd were listed. The order of the $k_{cat}/K_m$ values of BcelFp for ginsenosides (Rd > Rb1 > Rb2 > Rc) was the same as that observed for relative activity. However, the $V_{max}$ values and the substrate affinity of the enzyme followed the orders Rb1 > Rc > Rb2 > Rd, and Rd > Rb1 > Rb2 > Rc, respectively. BcelFp had higher catalytic efficiency for Rd rather than Rb1, Rb2 and Rc. The $k_{cat}/K_m$ value of BcelFp for ginsenoside Rd was 27.91 mM$^{-1}$s$^{-1}$, which was higher than that of β-glucosidase from G. terrae (2.21 mM$^{-1}$s$^{-1}$) (Shin et al. 2014), and β-glycosidase from P. mucilaginosus (1.78 mM$^{-1}$s$^{-1}$) (Cui et al. 2014).

**Table 4** Kinetic parameters for BcelFp with PPD-type ginsenosides

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_m$ ($\mu$M)</th>
<th>$V_{max}$ (mM/h)</th>
<th>$k_{cat}/K_m$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rb1</td>
<td>3.66±0.04</td>
<td>22.94±2.25</td>
<td>12.53</td>
</tr>
<tr>
<td>Rb2</td>
<td>4.02±0.12</td>
<td>18.87±1.56</td>
<td>9.39</td>
</tr>
<tr>
<td>Rc</td>
<td>5.95±0.03</td>
<td>22.67±1.79</td>
<td>7.62</td>
</tr>
<tr>
<td>Rd</td>
<td>0.67±0.006</td>
<td>9.35±1.37</td>
<td>27.91</td>
</tr>
</tbody>
</table>

**Discussion**

Cellulase has been extensively utilized in various industries including the agriculture, healthcare, food, bio-fuels and so on. The novel cellulose gene, termed *BcelFp*, was homologous to GH family 5 in the CAZy classification system. Presently, GH5, one of the largest GH families, exemplifies a family with a large variety of specificities. Until now, 29 known activities of GH5 glycoside hydrolases have been collected at the web of CAZy. These GH5 enzymes with different substrate specificities share the same retaining catalytic mechanism including one Glu residue as catalytic nucleophile/base and another Glu residue as catalytic proton donor. The alignment of BcelFp with three characterized glycoside hydrolases from GH5 family indicated that the Glu144 and Glu260 residues may be the catalytic residues in BcelFp. We can infer from these facts that the diversity in specificity of GH5 proteins may arise by the flexibility in
substrate binding sites. The broad specificity of GH5 enzymes makes it possible to further search more lots of different types of substrates and enlarge their application scope. The cellulase BcelFp, which exhibited ginsenoside-transforming activity, was a good example of exploring new activities of GH5 glycosidases. All ginsenoside-transforming glycosidases except for arabinofuranosidase and β-galactosidase from C. saccharolyticus (Shin et al. 2013) belong to GH 1 or 3 family. BcelFp was the first GH5 enzyme displaying the hydrolysis activity towards ginsenosides. Study of the cellulase BcelFp will inspire researchers to identify more and more ginsenoside-transforming glycosidases from GH5 family.

The cellulase BcelFp is active and stable under acidic conditions. It exhibited optimal activity at pH 6.0 similar to ginsenoside-transforming glycosidases from thermophilic origins such as β-glucosidases from Sulfolobus solfataricus (Noh et al. 2009), Pyrococcus furiosus DSMZ 3638 (Oh et al. 2014) and Thermotoga thermarum DSM 5069T (Pei et al. 2015). The optimal temperature of BcelFp was about 95 °C, which was higher than the cellulase from F. nodosum Rt17-B1(Wang et al. 2010). Even at 95 °C, more than 50% of the enzyme activity remained after incubation for 10 hours. The Tm of BcelFp was 96 °C, which was lower than β-glycosidase from Pyrococcus furiosus (Oh et al. 2014), but higher than the other thermophilic glycosidases. Thermostable enzymes serve as ideal catalysts for biotransformation application because high temperatures improve the ginsenosides solubility, enhance the substrate conversion and reduce the need for expensive cooling process. For instance, β-glucosidase from Thermotoga thermarum (Pei et al. 2015) transformed ginsenoside Rb1 into Rg3 with a corresponding molar conversion of 97.8% within 60 min at 85 °C. The β-glucosidase from Thermus thermophilus (Shin et al. 2014) transformed GypXVII to ginsenoside F2 with a molar yield of 100% at 90 °C.

The PPD-type ginsenosides including Rb1, Rb2, Rc and Rd can be easily extracted from ginseng roots. As seen in Table 3, Rb1, Rb2, Rc and Rd possess the same sugar moieties linked to C3 position in aglycon PPD. Both the C3 inner and outer sugars are glucose. The structural differences of these four ginsenosides lie in the sugar moieties linked to C20 position. The C20 inner sugar is glucose and the C20 outer sugar is glucose, arabinopyranose, xylose or arabinofuranose. Biotranformation of these four ginsenosides by glycosidases have been attempted previously. These enzymes can be divided into three groups according to the hydrolysis activity on the linked positions of sugar moieties in ginsenosides. Group I enzymes catalyzed the simultaneous hydrolysis of C3 and C20 sugars in Rb1, Rb2, Rc or Rd. For instance, β-glucosidase from Actinosynnema mirum exhibited the hydrolyzing activities as follows: Rb2→CO→CY, Rc→CMc1→CMc and Rd→F2→Rh2→PPD, respectively (Cui et al. 2012). Group II glycosidases catalyze the hydrolysis of C20 sugar moieties in Rb1, Rb2, Rc or Rd. For instance, β-glucosidase from M. esteraromaticum hydrolyzed the outer and inner glucoses attached to the C20 position of Rb1 along the following pathway: Rb1→Rd→Rg3 (Quan et al. 2011). Group III enzymes catalyze the hydrolysis of C3 sugars in Rb1, Rb2, Rc or Rd. Table 5 presented a summary concerning the recombinant glucosidases acting on the C3 sugars in Rb1, Rb2, Rc or Rd. Comparing Group I and II, very few Group III enzymes were identified. The cellulase BcelFp was proven to be Group III enzyme, as it cleaved the outer glucose moiety at the C3 carbon of ginsenoside Rb1, Rb2, Rc and Rd. The hydrolysis behavior of BcelFp on ginsenosides was similar to the GH1 β-glucosidase from S. alaskensis (Shin et al.
2014) and GH1 β-glucosidase from *C. cellulans* sp. 21 (CcBgl1A) (Yuan et al. 2015). However, BcelFp, CcBgl1A and the *S. alaskensis* glucosidase possessed different physicochemical properties, especially optimal temperature and substrate specificity. BcelFp was thermophilic cellulase, while CcBgl1A and the *S. alaskensis* glucosidase were mesophilic β-glucosidase. Moreover, the catalytic efficiency for hydrolysis of ginsenoside Rd by BcelFp was higher than that of *S. alaskensis* glucosidase (4.8 mM$^{-1}$min$^{-1}$) and CcBgl1A (8.64 mM$^{-1}$S$^{-1}$) from *C. cellulans* sp. 21. Hence, BcelFp with good thermostability and high catalytic efficiency was a promising biocatalyst for hydrolysis of PPD-type ginsenosides specifically for biotransformation of Rd.

### Table 5

The recombinant β-glucosidases act on the C3 sugars in Rb1, Rb2, Rc or Rd.

<table>
<thead>
<tr>
<th>Organism</th>
<th>GH</th>
<th>Reaction conditions</th>
<th>Substrate</th>
<th>Product</th>
<th>Cleavage site</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. cellulans</em> sp. 21</td>
<td>1</td>
<td>pH 5.5, 35 °C</td>
<td>Rb1, Rb2, Rc, Rd</td>
<td>GypXVII, CO, CMc1, F2</td>
<td>C-3</td>
<td>Yuan et al. 2015</td>
</tr>
<tr>
<td><em>S. alaskensis</em></td>
<td>1</td>
<td>pH 5.5, 50 °C</td>
<td>Rb1, Rb2, Rc, Rd</td>
<td>GypXVII, CO, CMc1, F2</td>
<td>C-3</td>
<td>Shin et al. 2014</td>
</tr>
<tr>
<td><em>P. aculeatum</em></td>
<td>3</td>
<td>pH 4.5, 70 °C</td>
<td>Rb2, Rc, Rd</td>
<td>CO, CY, CMc1, CMc, F2, CK</td>
<td>C-3</td>
<td>Lee et al. 2013</td>
</tr>
<tr>
<td><em>B. breve</em></td>
<td>3</td>
<td>pH 5.0, 35 °C</td>
<td>Rd</td>
<td>CK, F2</td>
<td>C-3</td>
<td>Zhang et al. 2019</td>
</tr>
</tbody>
</table>

In a conclusion, in the current study, a novel GH5 cellulase BcelFp was successfully cloned and expressed in *Escherichia coli*. The recombinant BcelFp exhibited optimal activity at pH 6.0 and 95 °C and showed high thermostability. BcelFp displayed high specificity and catalytic efficiency for biotransformation of ginsenosides Rb1, Rb2, Rc and Rd. Further investigation must be performed on large scale production of GypXVII, CO, CMc1, and F2 by using BcelFp.

**Declarations**

**Acknowledgements**

The authors acknowledge funding from Scientific and Technological Development Program of Jilin province of China (grant number 20210204046YY).

**Authors’ contributions**
SSY designed the research; FZ conducted the biochemistry and molecular experiments, and wrote the manuscript; HXZ conducted the qualitatively and quantitative investigation of the transformation process by high-performance liquid chromatography/mass spectrometry/mass spectrometry (HPLC/MS/MS), data analysis; NW, PZ, and KLZ involved in data analysis, and co-wrote the manuscript. All authors approved the final version of the manuscript.

**Funding**

This work was financially supported by funding from Scientific and Technological Development Program of Jilin province of China (grant number 20210204046YY).

**Availability of data and materials**

The datasets supporting the conclusions of this article are included within the article and its Additional file 1.

**Ethics approval and consent to participate**

This article does not contain any studies with human participants or animals performed by any of the authors.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

Jilin Ginseng Academy, Changchun University of Chinese Medicine, Changchun 130117, China.

**References**


**Figures**

**Figure 1**

Multiple amino acid sequence alignment of BcelFp with several characterized glycoside hydrolases from GH5. The accession numbers of the aligned sequences are for the following organisms: ADD73709, endoglucanase FnCel5A from Fervidobacterium nodosum Rt17-B1; AAD36816, endoglucanase from Thermotoga maritima MSB8; AXU72614. endoglucanase from Clostridioides difficile. The accession numbers were indicated to the left of the amino acid sequences. Identical residues are indicated by a red background. Symbols: ↑ amino acids forming a catalytic residues.

**Figure 2**

Phylogenetic analysis of BcelFp, and other characterized glycoside hydrolases from GH5. The units at the bottom of the tree indicate numbers of substitution events.
Figure 3

(A) Effect of pH on enzyme activity. (B) Effect of pH on enzyme stability. The activities were determined by assays with CMC as substrate following incubation of the enzyme at pH 4 (■), pH 5 (●), pH 6 (▲) and pH 7 (◀) for the indicated times. (C) Effect of temperature on enzyme activity. (D) Effect of temperature on enzyme stability. The activities were determined by assays with CMC as substrate following incubation of the enzyme at 85°C (▲), 90°C (●), and 95°C (■) for the indicated times. (E) Differential Scanning Calorimetry (DSC) analysis of BcelFp. The enzyme was concentrated to 1.5 mg/ml in 50 mM PBS buffer (pH 6.0). The equilibrated enzyme was scanned from 35 to 120 °C at a rate of 2.0 K/min. The enzyme scan was corrected using a buffer–buffer baseline.
Figure 4

HPLC analysis of ginsenoside Rb1, Rb2, Rc and Rd during biotransformation process by using BcElFp. Ginsenoside standards were indicated on the peaks. Numbers were used to indicate the product peaks.

Figure 5

Biotransformation pathways of ginsenoside Rb1, Rb2, Rc and Rd by using BcElFp.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.
• FigS1.jpg
• FigS2.tif
• FigS3.tif
• TableS1.docx