

RESEARCH

Open Access

Expression of an extremely acidic β -1,4-glucanase from thermoacidophilic *Alicyclobacillus* sp. A4 in *Pichia pastoris* is improved by truncating the gene sequence

Yingguo Bai^{1,2}, Jianshe Wang¹, Zhifang Zhang², Pengjun Shi¹, Huiying Luo¹, Huoqing Huang¹, Chunliang Luo¹ and Bin Yao^{*1}

Abstract

Background: *Alicyclobacillus* sp. A4 is thermoacidophilic and produces many glycoside hydrolases. An extremely acidic β -1,4-glucanase (Cela4) has been isolated from *Alicyclobacillus* sp. A4 and purified. This glucanase with a molecular mass of 48.6 kDa decreases the viscosity of barley-soybean feed under simulated gastric conditions. Therefore, it has the potential to improve the nutrient bioavailability of pig feed. For the study reported herein, the full-length gene, *CelA4*, of this glucanase (Cela4) was identified using the sequences of six peptides and cloned from strain A4. The gene fragment (*CelA4_F*) encoding the mature protein was expressed in *Pichia pastoris*. Sequence truncation and glycosylation were found for recombinant Cela4_F, both of which affected the expression efficiency. The physical properties of various forms of Cela4 as they affected enzymatic activity were characterized.

Results: We located the full-length 2,148-bp gene for Cela4 (*CelA4*) in the genome of *Alicyclobacillus* sp. A4. *CelA4* encodes a 715-residue polypeptide with a calculated molecular mass of 71.64 kDa, including an N-terminal signal peptide (residues 1-39), a catalytic domain (residues 39-497), and a C-terminal threonine-rich region (residues 498-715). Its deduced amino acid sequence and that of an *Alicyclobacillus acidocaldarius* endo- β -1,4-glucanase were identical at 44% of the residue positions. When the experimental molecular mass of Cela4_F--a recombinant protein designed to mimic the Cela4 sequence lacking the N-terminal signal peptide that had been expressed in *Pichia pastoris*--was compared with its hypothetical molecular mass, it was apparent that Cela4_F was truncated, possibly at residue 497. An artificially truncated gene fragment (*CelA4_T*) without C-terminal threonine-rich region was expressed in *P. pastoris*, and the expression efficiency of Cela4_T was substantially greater than that of Cela4_F. Purified Cela4_F and Cela4_T had similar molecular masses (~60 kDa) and enzymatic properties (optimum pH, 3.4; optimum temperature, 60°C); they were relatively stable between pH 1.2 and 8.2 at 70°C and resistant to acidic and neutral proteases. However, their molecular masses and thermostabilities differed from those of Cela4 isolated from *Alicyclobacillus* sp. A4. A deglycosylated form of Cela4 (Cela4_D) had properties similar to that of Cela4 except that it was thermolabile at 60°C.

Conclusions: Truncation during expression of Cela4_F or artificial truncation of its gene--both of which produced a form of Cela4 lacking a threonine-rich region that includes a putative linker--increased the level of enzyme produced in comparison with that produced by cultivation of *Alicyclobacillus* sp. A4. Glycosylation increased the thermostability of Cela4. Of the four forms of Cela4 studied, Cela4_T was produced in highest yield and had the most favorable physical properties; therefore, it has potential for use in the feed industry.

Background

β -Glucan is the major cell-wall component of cereals such as barley, wheat, oat, and rye [1], and it can be

hydrolyzed by β -glucanases. Microbial glucanases are often used in industry, including those related to waste management [2], alcohol fermentation [3], and animal feed production [4]. Several β -glucanases from the genus *Alicyclobacillus* have been identified, including two endo-glucanases (Cela and CelB) [5,6], one β -1,4-glucanase

* Correspondence: yaobin@caas-bio.net.cn

¹ Key Laboratory for Feed Biotechnology of the Ministry of Agriculture, Feed Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, PR China
Full list of author information is available at the end of the article

(CelA4) [7], one β -1,3(4)-glucanase (Agl9A) [8], and one cellulase (CelG) [9]. All are very thermoacidophilic and have optimum activities between 45 and 60°C and pH 2.0 and 6.0.

According to the primary structures of their catalytic domains, β -1,4-glucanases have been classified as members of the glycoside hydrolase (GH) families 5, 6, 7, 8, 9, 12, 44, 45, 48, and 51 http://www.cazy.org/fam/acc_GH.html. Most of the GH 51-type glucanases are α -L-arabinofuranosidases. Only three β -glucanases, endoglucanase F precursor from *Fibrobacter succinogenes* (AAC45377) [10], endoglucanase CelB from *A. acidocaldarius* (CAD86595) [6], and a cellulase from an uncultured bacterium (CAF22222.1) [11], are GH 51-type glucanases.

We previously purified the extremely acidic GH 51-type β -1,4-glucanase, CelA4 with a molecular mass of 48.6 kDa, from the thermoacidophilic *Alicyclobacillus* sp. A4 [7]. The pH optimum of CelA4 is 2.6, it is protease resistant, and can decrease the viscosity of barley-soybean feed under simulated gastric conditions. These properties indicate that CelA4 may improve the nutrient bioavailability of pig feed. For the commercialization of CelA4, recombinant gene expression in a high-throughput fermentation system is necessary. The methylotrophic yeast, *Pichia pastoris*, is an excellent host for the heterologous expression of recombinant proteins for which expression is controlled by the alcohol oxidase 1 promoter [12]. High-cell-density fed-batch cultivation usually consists of four phases and has been widely used to improve protein expression in *P. pastoris* [13]. The purpose of the study reported herein was to obtain the gene for CelA4 and, using it, to develop a high-yield fermentation process for CelA4 in *P. pastoris*. Upon doing so, we then examined how the physical properties of the native and recombinant enzymes affected enzymatic activity and identified certain properties that affected expression efficiency.

Results

Identification and sequence analysis of the full-length β -1,4-glucanase gene (CelA4)

We located the 2,148-bp full-length β -1,4-glucanase gene (Figure 1) (*CelA4*) in the *Alicyclobacillus* sp. A4 genome using six known peptide sequences of CelA4 as identifiers [7]. *CelA4* encodes a 715-residue polypeptide (calculated molecular mass, 71.64 kDa), which includes an N-terminal signal peptide (residues 1-39), a catalytic domain (residues 39-497), and a C-terminal threonine-rich region (residues 498-715, 21.12% threonine). Pro, Asp, Ser, and Glu--typical linker amino acids [14]--comprise 62% of residues 498-523, and therefore this region is presumed to be a linker. The enzyme also contains nine putative N-glycosylation sites having the consensus sequence, Asn-

Xaa-Thr/Ser-Zaa, where Zaa is not Pro; five of these sequences are in the catalytic region, and four are in the threonine-rich region. The deduced amino acid sequence of *CelA4* is most similar (44% identical) to that of the GH 51 cellulase, CelB, from *A. acidocaldarius* (CAD86595) [6]. The threonine-rich regions of these two enzymes have only 28% of their residues in common, and the threonine-rich region of CelA4 has <15% sequence identity with those of other glucanases. Alignment of CelA4 with five other glucanases using ClustalW is shown in Figure 2. The sequence alignment indicated that CelA4 does not contain a carbohydrate (cellulose)-binding domain found in the four glucanase sequences of ACU75486 (residues 625-724), EEP70239 (residues 600-699), ACU35994 (residues 575-676), and EEW74700 (residues 613-716). The putative catalytic residue in CelA4, Glu176, is highly conserved in glucanases and is located within the active site as predicted by sequence alignment [15,16].

Expression and purification of recombinant CelA4_F in *P. pastoris*

The gene (*CelA4_F*), which encodes a form of CelA4 that lacks the N-terminal signal sequence, was cloned into the pPIC9 vector that was then transformed into *P. pastoris* competent cells. The cells were cultured and clones were isolated. The clone that had the highest β -1,4-glucanase activity after flask cultivation was selected for expression in a 3.7-L fermenter. During fermentation, the cell mass of recombinant *P. pastoris* kept increasing from phase one (~110 g/L; about 18 hours) to phase four (~350 g/L; about 184 hours) (Figure 3) (See Methods for descriptions of the cultivation phases). The β -1,4-glucanase activity in the supernatant was 268 U/mL 156 h after induction with methanol.

Recombinant CelA4_F was purified by anion exchange chromatography. It migrated as a single band upon SDS-PAGE and had an apparent molecular mass of ~60 kDa (Figure 4), which is less than the predicted molecular mass (71.64 kDa) of CelA4_F, but greater than that of CelA4 (48.6 kDa). Using LC-ESI-MS/MS, this band was identified as CelA4_F. These results indicate that recombinant CelA4_F was probably both truncated and glycosylated.

Deglycosylation and artificial truncation of CelA4_F

Because CelA4 has nine putative N-glycosylation sites, the observed variation in apparent molecular mass of CelA4_F might be ascribed to glycosylation and/or truncation. After endo- β -N-acetylglucosaminidase H (Endo H) treatment, the protein migrated as a single band of 48 kDa upon SDS-PAGE, and thus the decrease in apparent molecular mass was ascribed to deglycosylation. The molecular mass of deglycosylated CelA4_F (CelA4_D) is

```
1      ATGTCACCTTCAGGGGGAGTCTGTGTGAACCGAAAACAGCGTACGTTAAAGTTGGGAACGCTCGCAGCAACGATTGTGGCACTCTCAGCC
1      M S P S G G V C V N R K Q R T L K L G T L A A T I V A L S A
91     GTGGCCACGCCTGCGGTGCGCCAGTGCAGGATACAGCAGCGCCATTGCGTCAICGACAGTTCATGTCACAGTCAATGCCGCTGCTGAACTT
31     V A T P A V A S A D T T T A I A S S T V H V T V N A A A E L
181    GGAATCGTGCCCAATACTGCACCTTGGTGTGAATACGGCCGCTCTGGGACGGGCATTACTCGATGCAGCCATTCCATCTCTGCTTCGTGGC
61     G I V P N T A L G V N T A V W D G H L L D A A I P S L L R G
271    ATTGGGGTAACCATGTTGCGATATCCCGGAGGATCGACTTCAGATGAGTACAATTGGCACAACGAATACCGTAACCTGGGGGTTATGCAGAT
91     I G V T M L R Y P G G S T S D E Y N W Q T N T V T G G Y A D
361    CCCAACACACCTTTGACAACTTCATGGGAGTGGTCCAAAAGGCTGGTGCAGCAACCCATTATTACGGTCAACGCCGGCACGGGCACACCG
121    P N N T F D N F M G V V Q K A G A Q P I I T V N A G T G T P
451    AGTGAAGCTGCCGCATGGGTTCAAGATGCAAAATGTCAACGCACCACTACGGTGTCAAGTATTGGGAAATCGGAAATGAGATGTATGGCAGC
151    S E A A A W V Q D A N V T H H Y G V K Y W E I G N E M Y G S
541    TGGGAAGCAGGGAATTTTGCAATAACCCATCTGGTTATGCGAAAGAGCTGTATCCTTCATTACGGCCATGAAAGCGGTTGATCCTTCT
181    W E A G N F A N N P S G Y A K E A V S F I Q A M K A V D P S
631    ATTAARATCGGCGTGGACCTCATCGCACCTGGTACTGGAGAAGATGACTGGAAATGCAACCGTGCCTTAGCACCAIGCACAGCCTTGGGGTG
211    I K I G V D L I A P G T G E D D W N A T V L S T M H S L G V
721    CTTCCGGACTTCGCCATTGTGCACCTGGTATGCGCAAAATCCGGGAGGCGAGACGGACGCTGGGCTGCTCAGTTCGACGAATCAAATTTG
241    L P D F A I V H W Y A Q N P G G E T D A G L L S S T N Q I S
811    ACGATGATGGATACTTGAAGCAGCAATTGAGCTCCTATGGAACCATCCCGGTTTTTCGTCACCGAAACGAATTCGGTTTTCGTATAACCC
271    T M M D T L K Q Q L S S Y G T I P V F V T E T N S V S Y N P
901    GGGCGCCAGAGTACAAGTCTGGTTAATGCACCTGTTCCCTCGATGATGACATGGCAGACTGGCTTGAGTCCGGAGCGCAAAACGTCGACTGG
301    G R Q S T S L V N A L F L D D D M A D W L E S G A Q N V D W
991    TGGGACTTGCCATAACGGCATTGTACACAACAGGCGGTGCAAAATGTGATCCGAATCTGTATGGGAGTATAACTACGGAGACTATGGA
331    W D L H N G I V T Q Q A G A N V D P N L Y G Q Y N Y G D Y G
1081   CTTTTGTCCAAATGGCTCGAGTGCACATGGCATTTCAGAACCGGCTGCCAATACACCATTCCCAACGTATTATGGATAACCAATGCTCGCA
361   L L S N G S S D N G I S E P A A N T P F P T Y Y G Y Q M L A
1171   GCCGTCATGGTGCCAGGAGCGACGATGATCGGTGCGGATCGAACAATGATCTGGTGGCCGTCATCGCACCAAGTTGCCAATGGTGTCT
391   A V M V P G A T M I G A G S N N D L V A V H A T K L P N G A
1261   GTCGACGTCATGTTGATCAACAAGGATCCGAAACAAGCGTATACCGTCGATTGCAAGCTGAAGGATTTGCTGCTAAGGGTCCCTGCATTC
421   V D V M L I N K D P K Q A Y T V D L Q A E G F A A K G P A F
1351   ACGTTATTCTACGGGCAAGGCAGTAACGCGGTGACCGCGGGCAAAATGGATAATCTGCAGAAATGTGACACTACCGCCCTATTCTGTGACG
451   T L F Y G Q G S N A V T P G K L D N L Q N V T L P P Y S V T
1441   GACATCATACACCAGCGGTGCCCGGGATCAGCCACAAGGGCCACAGTTTACGGACAAGACGAGTTATCCACTCCTCAGGTAAGGCT
481   D I I I P A V P G H Q P Q G P Q F T D K T T L S T P Q V K P
-----
1531   AGTGCCACAGAGACTTTGACCACGACGTTTACCAGACCGCGTGGTGGGTCAGGATGGTACACTCGACGTGGAATCTACAATCCAGCA
511   S A N E T L I T T F T D I R G A V K D G T L D V E I Y N P A
-----
1621   GGGCAATGGTTGGGCAACAAGTGCAGTCTGGCGTGACGTTTACGCTGGGCAATCATCTCAACCGATTACCTGGAACCTGGACGGCGGCC
541   G Q L V G Q Q V Q S G V T F T P G Q S S Q P I T W N W T A P
-----
1711   GATTCTCCTGGGACGTATAACCGTGAAGGCGTTCGTTCTCAGCCAAGACGGAACAAGCGTGTATGCGGCAGACCCGAGTGCAGCTACGTT
571   D S P G T Y T V K A F V F S Q D G T S V Y A A D P S A A T F
-----
1801   ACGGTCACACAGCCGGATCCGCCACCAATTCGGCCACCGTTACGCTGTCCGCAACTACGGTCAAAGTGGGTACACCTGTGACCATCAGC
601   T V T Q P D P P T I S A T V Q L S A T T V K V G T P V T I T
-----
1891   ACGACTTACACCGAAACCGCGCTACGGGGTACCTGAACAACGGGTTGCTTGTACAGTACCGCGTACAACTGACATCATCGCAA
631   T T Y T E T A P T G Y L N N G L L V Q Y A V Y N N W T S S Q
-----
1981   CAGTCCAATCCAACCTGCGACATTGACTCCTGGGCAATCGGTGACTGAGACTTGGACATTTACGCCAGAGCAGGCGGAACTACACATTC
661   Q S N P T A T L T P G Q S V T E T W T F T P E Q A G T Y T F
-----
2071   CCTGAAGGCATCTTACCAGTGGATGGACACAATTGCAGTGGATTAATCAGAACGTCACCTTACTGTGACAACTAA
691   P E G I F T S G W T Q L Q W I N Q N V T L T V T N *
```

Figure 1 Nucleotide and amino acid sequences of the β -1,4-glucanase, CelA4, from the thermoacidophilic *Alicyclobacillus* sp. A4. The residues of the putative signal peptide (residues 1-39) are underlined, those of the linker region are underlined twice, and the threonine-rich region is indicated by dashed line, nine putative N-glycosylation sites are indicated by black boxes.

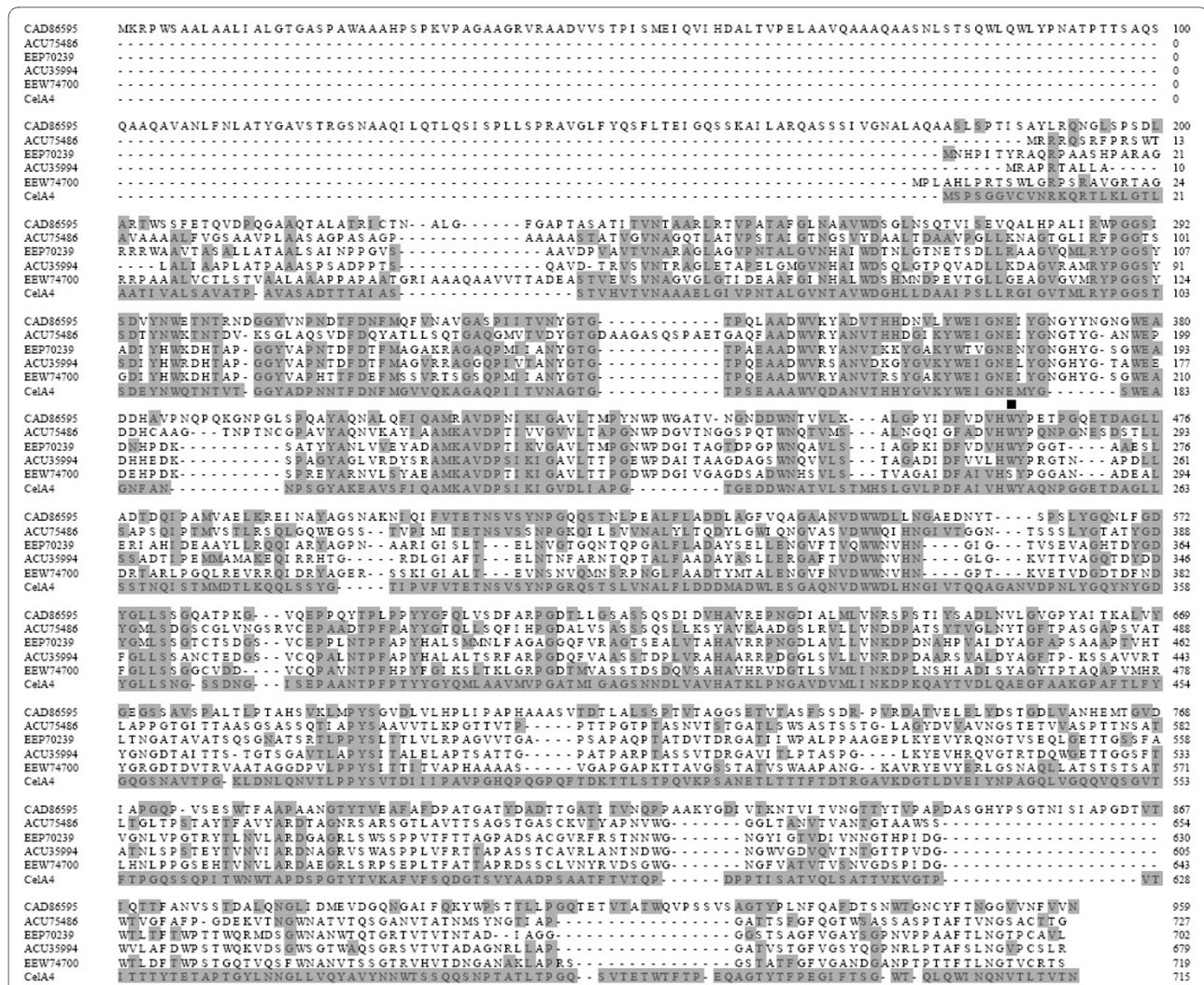


Figure 2 Amino acid sequence alignment of CelA4_F and five other glucanases. The sequences were those of glucanases from *Alicyclobacillus acidocaldarius* DSM 446 (CAD86595), *Catenulispora acidiphila* DSM 44928 (ACU75486), *Micromonospora* sp. ATCC 39149 (EEP70239), *Actinosynnema mirum* DSM 43827 (ACU35994), and *Streptomyces flavogriseus* ATCC 33331 (EEW74700). ClustalW was used to align the sequences. Residues that are the same at all positions are indicated by solid grey boxes; the catalytic residue, Glu176, is indicated by a solid square.

almost the same as that of CelA4. Therefore, CelA4_T was both glycosylated and truncated when expressed in *P. pastoris*.

According to the molecular mass comparison and sequence analysis, residue 497 was predicted to be the truncation site. CelA4_T was expressed in *P. pastoris* as described above. The *P. pastoris* culture for CelA4_T followed the same growth profile as for CelA4_F. The β-1,4-glucanase activity in the supernatant of the CelA4_T culture reached 420 U/mL (Figure 3), which was much higher than that found for the CelA4_F culture. Purified CelA4_T migrated as a single band upon SDS-PAGE and had a molecular mass of ~60 kDa, which is the same as found for CelA4_F.

Enzyme characterization

The physical properties of truncated CelA4_T, artificially truncated CelA4_D, deglycosylated and truncated CelA4_D, and CelA4 [7] that affect enzyme activity were characterized and compared. CelA4_F, CelA4_T, and CelA4_D showed optimum activity at pH 3.4, as opposed to pH 2.6 for native CelA4. The first three enzymes were less active between pH 1.2 and 2.2 and between 4.8 and 7.6 than was native CelA4 (Figure 5A). CelA4_F, CelA4_T, CelA4_D, and CelA4 were all stable between pH 1.8 and 8.2, at 37°C for 1 h, but the first three enzymes were more stable at pH 1.2 (>70% retention of activity) than was CelA4 (~36% retention of activity) (Figure 5B). All of the enzymes displayed maximum activity between 60 and 65°C (Figure 5C). Notably, CelA4_F and CelA4_T were relatively stable at

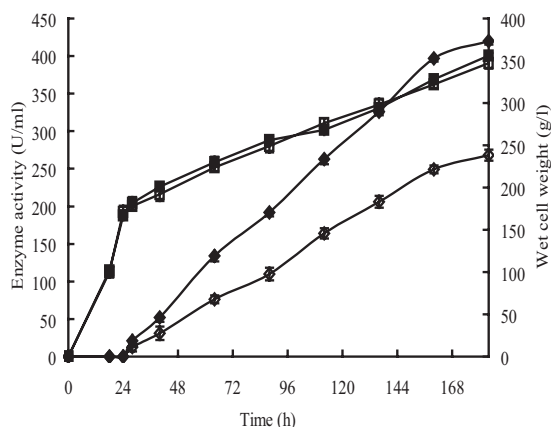


Figure 3 Time courses for the appearance of β -1,4-glucanase activity and increase in biomass during fermentation. The symbols and associated proteins are as follows: diamond stands for activity; Square stands for biomass; hollow stands for CelA4_F and solid stands for CelA4_T.

75°C, as they maintained > 75% of their initial activities after a 1-h incubation at pH 3.4, 75°C. Under the same conditions, CelA4_D and CelA4 lost all activity (Figure 5D).

CelA4_F, CelA4_T, and CelA4_D were highly resistant to the acidic and neutral proteases, including trypsin, α -chymotrypsin, collagenase, pepsin, and subtilisin A, and they retained more than 70% of their activities after incubation with these proteases at 37°C for 2 h. This resistance to

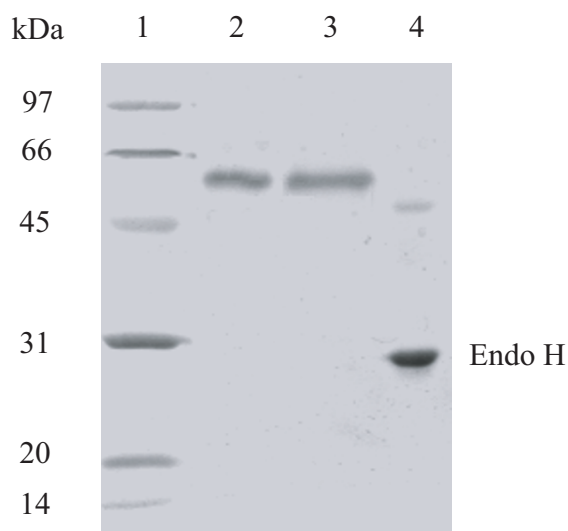


Figure 4 SDS-PAGE gels of purified CelA4_F and CelA4_T before and after Endo H treatment. Lanes: 1, standard protein molecular-weight markers; 2, purified CelA4_F; 3, purified CelA4_T; 4, purified CelA4_F after deglycosylation with Endo H.

acidic and neutral proteolysis had been found previously for CelA4 [7].

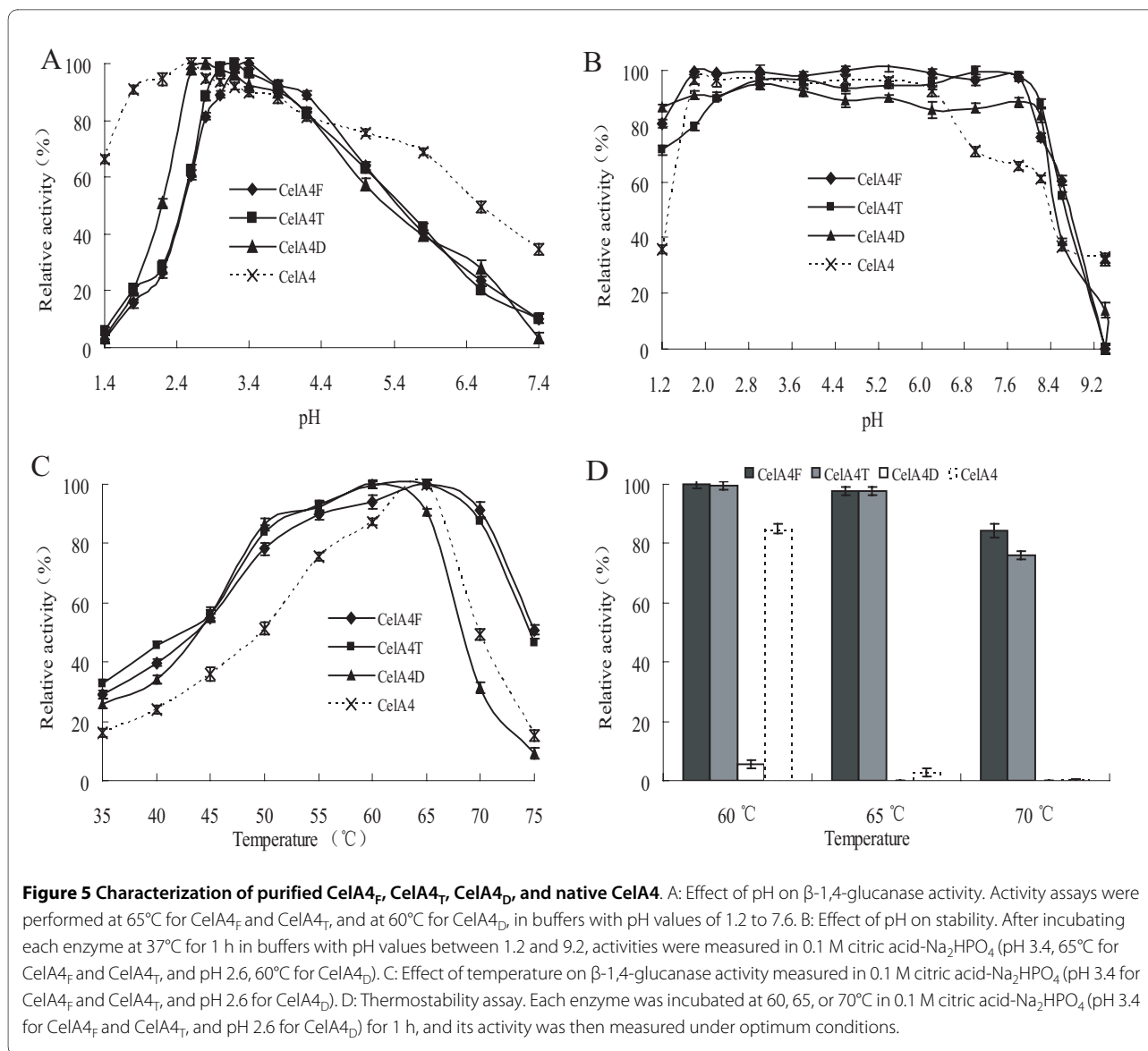
Nucleotide sequence accession number

The nucleotide sequence of the β -1,4-glucanase gene (CelA4) from *Alicyclobacillus* sp. A4 was deposited in GenBank under the accession number [GU576556](https://www.ncbi.nlm.nih.gov/nuclot/GU576556).

Discussion

We previously isolated an extremely acidic GH 51 β -glucanase, CelA4, from thermoacidophilic *Alicyclobacillus* sp. A4. Herein, we described the gene sequence and the expression of CelA4_P, which encodes CelA4 lacking the N-terminal signal sequence. Its deduced amino acid sequence is only 42% identical to that of CelB from *A. acidocaldarius* [6], indicating that CelA4 is a previously uncharacterized gene. The deduced C-terminal 217-residue sequence of CelA4 is threonine rich (21.2%), which is a much higher than found on average (5.74%) for proteins [17]. Threonine-rich regions have been reported to be involved in fibronectin binding [18], unidirectional transport of a mineralocorticoid receptor into the nucleus [19], vanadate resistance [20], and resistance to HIV by binding to a specific receptor [21]. Structural analysis has indicated that the functions of threonine-rich regions are usually associated with those of linker regions and O-linked glycosylation [22,23]. We have found that the threonine-rich region containing the putative linker of native CelA4 and CelA4_F was removed during secretion from *Alicyclobacillus* sp. A4 and *P. pastoris*, respectively. Therefore, the threonine-rich region may act as a molecular chaperone and be involved in proper folding of the catalytic domain. The linker has been reported to be necessary for thermostability [24,25]. Because it is removed with the threonine-rich sequence, we assume that the linker has no effect on the catalytic properties of CelA4.

Native CelA4 has excellent properties and therefore has great potential for use in industrial applications [7]. However, the yield of CelA4 was very small (0.9 U/mL) when *Alicyclobacillus* sp. A4 was cultured in a glucanase-inducing medium. For the study reported herein, we used a 3.7-L fermenter for the cultivation of *P. pastoris* containing a plasmid carrying CelA4_F. After expression, 268 U/mL of CelA4 activity was measured in the culture supernatant, which is approximately 300-fold greater than found upon cultivation of *Alicyclobacillus* sp. A4. We also constructed a gene for the C-terminally truncated glucanase, CelA4_T that lacked the threonine-rich region using a PCR-based gene truncation method, expressed the protein in *P. pastoris*, and obtained a yield of 420 U/mL. Expression of truncated genes has been shown to increase enzyme production and improve both activity and thermostability [26,27]. Truncated CelA4_F and artificially truncated



Cella4_T have similar molecular masses and enzymatic properties but differ in the relative amount of enzyme produced by fermentation. This difference might be ascribed to the intracellular functions of the threonine-rich region. It is possible that the threonine-rich region acts as a molecular chaperone, but how it influences enzyme production is unknown.

Both Cella4_F and Cella4_T were glycosylated when expressed in *P. pastries*. It has been reported that glycosylation has significant effects on enzyme thermostability [28], the optimum pH value [29], and resistance to proteases [30,31]. Cella4_D, Cella4_F, and Cella4_T had similar pH and stability profiles, but Cella4_D was not as thermally stable as the other two proteins. For example, Cella4_D exhibited ~30% of its maximal activity at 70°C, whereas

both Cella4_F and Cella4_T retained ~90% of their initial activities. Moreover, Cella4_F and Cella4_T were thermostable at 70°C, but Cella4_D was not stable at 60°C. Although the molecular mass of Cella4_D was similar to that of native Cella4, these enzymes differed with respect to their pH and temperature profiles and, most notably, their stabilities. Therefore, heterologous expression of Cella4_F increased the amount of the enzyme produced and improved the thermostability by incorporating sugar residues post-translationally.

Cella4_T and Cella4_F exhibited similar enzymatic properties, namely a pH optimum of 3.4, a temperature optimum of 65°C, stability between pH 2.0 and 8.2 and at 70°C, and they were both active and stable under simulated gastric conditions. Both enzymes could decrease

the viscosity of barley-soybean feed (data not shown), but more of the former was produced during *P. pastoris* fermentation. Therefore, in the future we will produce CelA4_T, not CelA4_F, by heterologous expression for commercial applications.

Conclusions

For the study reported herein, we identified and cloned the β -1,4-glucanase-encoding gene (*CelA4*) found in *Alicyclobacillus* sp. A4, and achieved high-yield expression of CelA4_F in *P. pastoris*. CelA4_F was truncated and glycosylated during fermentation. Expression of CelA4_T, which lacks the threonine-rich region, produced greater amounts of protein and had glucanase activity identical to that of CelA4_F. We speculate that the threonine-rich region might act as a molecular chaperone that ensures proper folding of the catalytic domain. Glycosylation was necessary for the thermostability of both CelA4_F and CelA4_T. All of our data indicate that recombinant CelA4_T produced using a *P. pastoris* fermentation system will have great potential as a β -1,4-glucanase for use in the feed industry.

Methods

Strains, plasmids, and chemicals

The strain *Alicyclobacillus* sp. A4 was deposited in the China General Microbiological Culture Collection Center under the registration number CGMCC3147 [32]. *Escherichia coli* JM109 was obtained from TaKaRa (Dalian, China). *P. pastoris* GS115 and the pPIC9 vector were obtained from Invitrogen (San Diego, CA, USA). Barley β -glucan was supplied by Sigma (St. Louis, MO, USA). T4 DNA ligase and restriction endonucleases were obtained from Promega (Madison, WI, USA). All other chemicals were of analytical grade and commercially available.

Cloning and expression of the β -1,4-glucanase gene (CelA4_F)

Using the sequences of an N-terminal peptide and five internal peptides of native CelA4 [7], the full-length coding gene for CelA4 (*CelA4*) was identified in the genome of *Alicyclobacillus* sp. A4 using BLASTp (Sequencing of the complete *Alicyclobacillus* sp. A4 genome is in progress). The sequence of the N-terminal signal peptide was predicted using SignalP <http://www.cbs.dtu.dk/services/SignalP/>. Alignment of multiple protein sequences was accomplished using ClustalW [33]. Vector NTI 10.0 software was used to identify homologous and identical residues after sequence alignment and to predict the molecular mass of the mature protein.

To construct the plasmid containing *CelA4_F*, the gene encoding CelA4_F was PCR-amplified using the primers

CelA4_FF: GCATACGTAATGGAGGCGACTATGCAA GCAGC and CelA4_FR: GAAGCGGCCGCTCAGACACCACAAAATGAGAAACCAC (*Sna*BI and *Not*I sites are bold) with *Alicyclobacillus* sp. A4 genomic DNA as the template. The PCR-amplified gene fragment was cloned in-frame at the downstream site of the α -factor (signal peptide) present in the pPIC9 vector and transformed into *P. pastoris* competent cells. Positive transformants were cultured in minimal dextrose medium or minimal methanol medium, and the culture supernatants were screened for glucanase activity.

High-cell-density fermentation of recombinant CelA4_F

The positive transformant with the highest level of β -1,4-glucanase activity was grown in a 3.7-L fermenter (Bioengineering KLF 2000, Wald, Switzerland) using a four-step fermentation strategy [13] that was scaled up. The *Pichia* fermentation was performed according to instructions obtained from Invitrogen. Fermentation began with a batch growth phase in 2.0-L basic sodium medium and the following conditions were used: agitation speed, 1,000 rpm; ventilation rate, 1.6 vvm; temperature, 30°C; pH, 4.5 (adjusted with 6 M NH₄OH). The first phase was terminated when the glycerol in the basic sodium medium was consumed completely (about 18 h) and the cell mass had increased to > 100 g/L (wet cell weight). The second glycerol (50%, w/v) fed-batch phase was initiated at 70 mL/h for ~6 h to further increase the cell mass to ~170 g/L. During the third phase, a glycerol/methanol (8:1, v:v) mixture was fed at 25 mL/h for ~4 h to transition the culture from glycerol to methanol. In the fourth phase, methanol was fed at 6-7 mL/h for ~156 h. During the fermentation process, dissolved oxygen was kept above 20%. Culture samples were collected every 12 h and subjected to cell mass and enzyme activity analyses.

Purification and identification of recombinant CelA4_F

The culture supernatant (about 2,500 mL) was concentrated ~15-fold by ultrafiltration using hollow-fiber membranes with molecular weight cut-offs of first 6 kDa (Motimo, Tianjin, China) and then 5 kDa (Vivascience, Hannover, Germany). The resulting solution was dialyzed against 20 mM McIlvaine buffer (0.2 M Na₂HPO₄ containing 0.1 M citric acid), pH 7.5, and loaded onto a HiTrap Q Sepharose XL FPLC column (Amersham Pharmacia Biotech, Uppsala, Sweden) equilibrated with the same buffer. Proteins were eluted using a linear gradient of NaCl (0-1.0 M) in the same buffer. Fractions with enzyme activity were collected and subjected to SDS-PAGE [34]. To identify the purified protein as CelA4_F, the protein band was excised from the gel, digested with trypsin, and sequenced using liquid chromatography/electrospray ionization tandem mass spectrometry

(MALDI-TOF-MS/MS) at the Institute of Zoology, Chinese Academy of Sciences.

Cloning of the artificially truncated β -1,4-glucanase gene (CelA4_T)

According to the molecular mass comparison of CelA4_F and native CelA4 and the sequence analysis, residues 498-715 was supposed to be hit by profiles with a high probability of occurrence (Entry: PS50325) based on ScanProsite analysis <http://www.expasy.ch/tools/scanprosite>. Residue 497 was predicted to be the truncation site. CelA4_T-encoding the truncated protein (residues 40-497) was PCR-amplified using the primers CelA4_TF (GCATACGTAATGGAGGCGACTATGCAAGCAGC) and CelA4_TR (GAAGCGGCCGCAAAGTGTGGCCCT-TGTGGCTGATGC) (*Sna*BI and *Not*I sites are bold) with genomic DNA of *Alicyclobacillus* sp. A4 as the template. CelA4_T was expressed in *P. pastoris* and subjected to the same purification procedure as described above.

Deglycosylation of CelA4_F

Purified CelA4_F (a ~2 μ g) was treated with 20 U of Endo H for 2 h at 37°C according to the supplier's instructions (New England Biolabs, Ipswich, MA, USA) and then analyzed by SDS-PAGE.

Enzyme assay

All enzymatic assays were performed in triplicate. The final reaction systems contained 50 μ L of an appropriately diluted enzyme and a 450- μ L solution containing 1% barley β -glucan (w/v) and 200 mM McIlvaine buffer at the optimum pH previously determined for each enzyme. Reactions were allowed to proceed for 10 min at the optimum temperatures and then terminated by adding 1.5 mL dinitrosalicylic acid [35]. Each mixture was heated in a boiling water bath for 5 min. After addition of 1.0 mL H₂O, the absorbance of each mixture at 540 nm was measured. One unit of enzyme activity was defined as the amount of enzyme that catalyzed the formation of 1 μ mol glucose per minute.

Biochemical characterization of purified CelA4_F, CelA4_T, and CelA4_D

The pH optima for the enzyme activities of CelA4_F and CelA4_T were determined at 65°C and at 60°C for CelA4_D. The enzymes were incubated for 1 h at 37°C in the absence of substrate, and then their activities were measured at pH 3.4 and 65°C for CelA4_F and CelA4_T, and at pH 2.6 and 60°C for CelA4_D. The buffers used were 0.1 M KCl-HCl for pH 1.2-2.2, 0.1 M McIlvaine buffer for pH 2.6-7.6, 0.1 M Tris-HCl for pH 8.0-9.0, and 0.1 M glycine-NaOH for pH 9.4-12.0. The optimal temperature for enzyme activity was determined using the McIlvaine buf-

fer (pH 3.4 for CelA4_F and CelA4_T; pH 2.6 for CelA4_D) at temperatures between 35°C and 75°C. The thermostability of each enzyme was determined by incubating the enzyme (100 μ g/mL) in McIlvaine buffer (pH 3.4 for CelA4_F and CelA4_T; pH 2.6, for CelA4_D) at 60°C, 65°C, or 70°C without substrate for 1 h and then measuring the enzyme activity under optimum conditions.

To examine resistance to different proteases, purified CelA4_F, CelA4_T, and CelA4_D (2 μ g/mL) were each incubated with 200 μ g/mL trypsin, α -chymotrypsin, collagenase, pepsin, or 500 μ g/mL subtilisin A at 37°C and at the pH optimum of the specific protease for various lengths of time. Incubations in the absence of each protease served as the controls. Activities were determined under the standard assay conditions of pH 3.4, 65°C for CelA4_F and CelA4_T, and of pH 2.6, 60°C for CelA4_D.

Abbreviations

Endo H: endo- β -N-acetylglucosaminidase H; SDS-PAGE: sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YB participated in the design of the study, the fermentation development, data analysis, and in the writing of the manuscript. JW participated in the construction of the recombinant plasmids. CL participated in the fermentation procedures and data analysis. YF participated in the biochemical characterization of the enzymes. ZZ participated in the writing and reviewing of the manuscript. BY participated in the writing and editorial supervision of the manuscript. PS supervised the experiments. HH participated in the gene sequence analysis. HL participated in the result analysis and the writing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This research was supported by the National High Technology Research and Development Program of China (863 program, grant 2007AA100601) and the earmarked fund for modern agro-industry technology research system (NYCYTX-42-G2-05).

Author Details

¹Key Laboratory for Feed Biotechnology of the Ministry of Agriculture, Feed Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, PR China and ²Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, PR China

Received: 12 March 2010 Accepted: 14 May 2010

Published: 14 May 2010

References

1. Lee C, Horsley R, Mantney F, Schwarz P: **Comparisons of β -glucan content of barley and oat.** *Cereal Chemistry* 1997, **74**:571-575.
2. Liu B, Yang Q, Zhou Q, Song J, Chen D, Liu H: **Cloning and expression of the endo- β -glucanase III cDNA gene from *Trichoderma viride* AS3.3711.** *Acta Environment Sinica* 2004, **25**:127-132.
3. Celestino KR, Cunha RB, Felix CR: **Characterization of a β -glucanase produced by *Rhizopus microsporus* var. *microsporus*, and its potential for application in the brewing industry.** *BMC Biochemistry* 2006, **7**:23.
4. Walsh GA, Murphy RA, Killeen GF, Headon DR, Power RF: **Technical note: detection and quantification of supplemental fungal β -glucanase activity in animal feed.** *Journal of Animal Science* 1995, **73**:1074-1076.
5. Eckert K, Zielinski F, Lo Leggio L, Schneider E: **Gene cloning, sequencing, and characterization of a family 9 endoglucanase (CelA) with an unusual pattern of activity from the thermoacidophile *Alicyclobacillus***

- acidocaldarius* ATCC27009. *Applied Microbiology and Biotechnology* 2002, **60**:428-436.
6. Eckert K, Schneider E: **A thermoacidophilic endoglucanase (CelB) from *Alicyclobacillus acidocaldarius* displays high sequence similarity to arabinofuranosidases belonging to family 51 of glycoside hydrolases.** *European Journal of Biochemistry* 2003, **270**:3593-3602.
 7. Bai Y, Wang J, Zhang Z, Shi P, Luo H, Huang H, Feng Y, Yao B: **Extremely acidic β -1,4-glucanase, CelA4, from thermoacidophilic *Alicyclobacillus* sp. A4 with high protease resistance and potential as a pig feed additive.** *Journal of Agriculture and Food Chemistry* 2010, **58**:1970-1975.
 8. Bai Y, Wang J, Zhang Z, Shi P, Luo H, Huang H, Luo C, Yao B: **A novel family 9 β -1,3(4)-glucanase from thermoacidophilic *Alicyclobacillus* sp. A4 with potential applications in the brewing industry.** *Applied Microbiology and Biotechnology* 2010. DOI: 10.1007/s00253-010-2452-3
 9. Morana A, Esposito A, Maurelli L, Ruggiero G, Ionata E, Rossi M, La Cara F: **A novel thermoacidophilic cellulase from *Alicyclobacillus acidocaldarius*.** *Protein and Peptide Letter* 2008, **15**:1017-1021.
 10. Malburg SR, Malburg LM Jr, Liu T, Iyo AH, Forsberg CW: **Catalytic properties of the cellulose-binding endoglucanase F from *Fibrobacter succinogenes* S85.** *Applied Environment and Microbiology* 1997, **63**:2449-2453.
 11. Grant S, Sorokin DY, Grant WD, Jones BE, Heaphy S: **A phylogenetic analysis of Wadi el Natrun soda lake cellulase enrichment cultures and identification of cellulase genes from these cultures.** *Extremophiles* 2004, **8**:421-429.
 12. Sreerishna K, Brankamp RG, Kropp KE, Blankenship DT, Tsay JT, Smith PL, Wierschke JD, Subramaniam A, Birkenberger LA: **Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*.** *Gene* 1997, **190**:55-62.
 13. Plantz BA, Sinha J, Villarete L, Nickerson KW, Schlegel VL: ***Pichia pastoris* fermentation optimization: energy state and testing a growth-associated model.** *Applied Microbiology and Biotechnology* 2006, **72**:297-305.
 14. Wootton JC, Federhen S: **Analysis of compositionally biased regions in sequence databases.** *Methods in Enzymology* 1996, **266**:554-571.
 15. Shallom D, Belakhov V, Solomon D, Gilead-Gropper S, Baasov T, Shoham G, Shoham Y: **The identification of the acid-base catalyst of α -arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase.** *FEBS Letter* 2002, **514**:163-167.
 16. Gilkes NR, Henriessat B, Kilburn DG, Miller RC Jr, Warren RA: **Domains in microbial β -1,4-glycanases: sequence conservation, function, and enzyme families.** *Microbiological Reviews* 1991, **55**:303-315.
 17. Karlin S, Bucher P, Brendel V, Altschul SF: **Statistical methods and insights for protein and DNA sequences.** *Annual Review of Biophysics and Biophysical Chemistry* 1991, **20**:175-203.
 18. Rauceo JM, De Armond R, Otoo H, Kahn PC, Klotz SA, Gaur NK, Lipke PN: **Threonine-rich repeats increase fibronectin binding in the *Candida albicans* adhesin Als5p.** *Eukaryotic Cell* 2006, **5**:1664-1673.
 19. Walther RF, Atlas E, Carrigan A, Rouleau Y, Edgecombe A, Visentin L, Lamprecht C, Addicks GC, Hache RJ, Lefebvre YA: **A serine/threonine-rich motif is one of three nuclear localization signals that determine unidirectional transport of the mineralocorticoid receptor to the nucleus.** *Journal of Biological Chemistry* 2005, **280**:17549-17561.
 20. Nakamura T, Namba H, Ohmoto T, Liu Y, Hirata D, Miyakawa T: **Cloning and characterization of the *Saccharomyces cerevisiae* SVS1 gene which encodes a serine- and threonine-rich protein required for vanadate resistance.** *Gene* 1995, **165**:25-29.
 21. Cotellet N, Lohez M, Cotellet P, Henichart JP: **Conformational study of the threonine-rich C-terminal pentapeptide of peptide T.** *Biochemistry and Biophysics Research Communication* 1990, **171**:596-602.
 22. Tyshenko M, d'Anjou M, Davies P, Daugulis A, Walker V: **Challenges in the expression of disulfide bonded, threonine-rich antifreeze proteins in bacteria and yeast.** *Protein Expression and Purification* 2006, **47**:152-161.
 23. Kieliszewski MJ, Leykam JF, Lampert DT: **Structure of the Threonine-Rich Extensin from *Zea mays*.** *Plant Physiology* 1990, **92**:316-326.
 24. Dias FM, Goyal A, Gilbert HJ, Jose AMP, Ferreira LM, Fontes CM: **The N-terminal family 22 carbohydrate-binding module of xylanase 10B of *Clostridium thermocellum* is not a thermostabilizing domain.** *FEMS Microbiology Letters* 2004, **238**:71-78.
 25. Li N, Shi P, Wang Y, Wang Y, Luo H, Bai Y, Zhou Z, Yao B: **A xylanase with high pH stability from *Streptomyces* sp. S27 and its carbohydrate-binding module with/without linker-region-truncated versions.** *Applied Microbiology and Biotechnology* 2008, **83**:99-107.
 26. Paloheimo M, Mantyla A, Kallio J, Puranen T, Suominen P: **Increased production of xylanase by expression of a truncated version of the xyn11A gene from *Nonomuraea flexuosa* in *Trichoderma reesei*.** *Applied Environment and Microbiology* 2007, **73**:3215-3224.
 27. Wen TN, Chen JL, Lee SH, Yang NS, Shyur LF: **A truncated *Fibrobacter succinogenes* 1,3-1,4- β -D-glucanase with improved enzymatic activity and thermostolerance.** *Biochemistry* 2005, **44**:9197-9205.
 28. Clark SE, Muslin EH, Henson CA: **Effect of adding and removing N-glycosylation recognition sites on the thermostability of barley α -glucosidase.** *Protein Engineering, Design and Selection* 2004, **17**:245-249.
 29. Somera AF, Pereira MG, Souza Guimaraes LH, Polizeli Mde L, Terenzi HF, Melo Furriel RP, Jorge JA: **Effect of glycosylation on the biochemical properties of β -xylosidases from *Aspergillus versicolor*.** *Journal of Microbiology* 2009, **47**:270-276.
 30. Wyss M, Pasamontes L, Friedlein A, Remy R, Tessier M, Kronenberger A, Middendorf A, Lehmann M, Schnoebelen L, Rothlisberger U, et al.: **Biophysical characterization of fungal phytases (myo-inositol hexakisphosphate phosphohydrolases): molecular size, glycosylation pattern, and engineering of proteolytic resistance.** *Applied Environment and Microbiology* 1999, **65**:359-366.
 31. Takegawa K, Yoshikawa M, Mishima T, Nakoshi M, Iwahara S: **Determination of glycosylation sites using a protein sequencer and deglycosylation of native yeast invertase by endo- β -N-acetylglucosaminidase.** *Biochemistry International* 1991, **25**:585-592.
 32. Bai Y, Wang J, Zhang Z, Yang P, Shi P, Luo H, Meng K, Huang H, Yao B: **A new xylanase from thermoacidophilic *Alicyclobacillus* sp. A4 with broad-range pH activity and pH stability.** *Journal of Industrial Microbiology and Biotechnology* 2010, **37**:187-194.
 33. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Research* 1994, **22**:4673-4680.
 34. Laemmli UK: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature* 1970, **227**:680-685.
 35. Miller GL: **Use of dinitrosalicylic acid reagent for determination of reducing sugar.** *Analytical Chemistry* 1959, **31**:426-428.

doi: 10.1186/1475-2859-9-33

Cite this article as: Bai et al., Expression of an extremely acidic β -1,4-glucanase from thermoacidophilic *Alicyclobacillus* sp. A4 in *Pichia pastoris* is improved by truncating the gene sequence *Microbial Cell Factories* 2010, **9**:33

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

