

# Identification of a Novel Garlic Cellulase Gene

Aeri Kim · Ryong Nam Kim · Dae-Won Kim ·  
Sang-Haeng Choi · Aram Kang · Seong-Hyeuk Nam ·  
Hong-Seog Park

Published online: 7 January 2010

© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** Genes encoding cellulase enzymes have been investigated in various plants due to the importance of cellulase enzymes in industrial applications, especially in the conversion of biomass into biofuels. Although several cellulase genes have been cloned and characterized, little is known about cellulase genes from garlic or enzyme activities of their gene products. In this study, a cellulase gene from garlic was cloned and characterized in gene and protein levels for the first time. The DNA sequence of the garlic cellulase gene showed 81% identity with the sequence of the endo-beta-1,4-glucanase of *Pisum sativum*.

The open reading frame of this gene is 1,506 bp, which corresponds to 501 deduced amino acids. We identified the novel ORF region, which was translated into a 55.2 kDa protein using the protein expression vector, pET28a, in *Escherichia coli* and we confirmed that this protein has cellulase activity in vitro. Our study demonstrates that garlic is very useful, not only for the culinary and pharmaceutical industries, but also as an excellent natural source of various kinds of important genes and enzymes.

**Keywords** Garlic · Cellulase · Cloning · Protein expression

---

Aeri Kim and Ryong Nam Kim contributed equally to this work.

A. Kim · R. N. Kim · D.-W. Kim · S.-H. Choi · A. Kang ·  
S.-H. Nam · H.-S. Park (✉)  
Genome Research Center and Industrial biotechnology  
and bioenergy research center,  
Korea Research Institute of Bioscience and Biotechnology,  
Daejeon 305-306, South Korea  
e-mail: hspark@kribb.re.kr

A. Kim  
e-mail: aeri\_kim@ust.ac.kr

R. N. Kim  
e-mail: krn1968@kribb.re.kr

D.-W. Kim  
e-mail: todaewon@kribb.re.kr

S.-H. Choi  
e-mail: shchoi@kribb.re.kr

A. Kang  
e-mail: rami@ust.ac.kr

S.-H. Nam  
e-mail: odysseus@kribb.re.kr

A. Kim · A. Kang · S.-H. Nam · H.-S. Park  
University of Science and Technology (UST),  
Daejeon 303-333, South Korea

## Introduction

Cellulose is the most common organic compound in nature and structural component of the plants. The primary cell wall of green plants consists of a linear chain of several hundred to over 10,000  $\beta$ -1,4-glucosidic bonds linking D-glucose units. Plant cell walls can be degraded into glucose through synergistic cellulolysis by three types of cellulase: endo-beta-1,4-glucanase, cellobiohydrolase, and  $\beta$ -glucosidase (Krause et al. 2003). Glucose derived from cellulose can be used as a natural source of energy by ruminants. Interestingly, a large number of microbes has been found in the cecum of ruminants. The cecum is important for cellulose digestion, helping ruminants obtain energy from plants, because ruminants cannot make their own cellulase for the digestion of cellulose into glucose (Krause et al. 2003).

So far, cellulase activities have been reported in various kinds of plants, and they have been known to play roles in the softening of climacteric fruits, and the abscission of leaves, flowers, and fruits (Del and Bennett 1996; Ferrarese et al. 1995). This event, called cell separation, is part of a more general phenomenon that occurs as a consequence of

the cell wall weakening caused by the activity of one or more hydrolases, for instance, polygalacturonases and cellulases (Osborne et al. 1989). In fact, the cellulase is one of the most widely studied enzymes because it has been used in commercial food processing, the fermentation of biomass into biofuels, the textile industry, laundry detergents, and the pulp and paper industries.

Although previous studies reported that garlic had cellulase activity, the studies were based only upon the measurement of enzymatic activities in extracts obtained directly from garlic juice, and the observation that the enzymatic activities are dependent upon pH or on the concentration of NaCl (Chang et al. 1993). The studies did not reveal the molecular basis of the cellulase activity.

The goal of this study was to clone the novel cellulase gene of garlic, to identify the protein product of this gene, and to assess the protein's cellulase activity. Our study is the first report to characterize the cellulase gene of garlic at the molecular level.

## Materials and Methods

### RNA Isolation and cDNA Library Construction

The garlic used in this study, the Korean Nam-do cultivar, was purchased from the Chungnam National Institute of Agricultural Science and Technology. Total RNA was isolated from stem tissue of the garlic plant by the same method used for pine tree RNA (Juan et al. 2001). A cDNA library containing 5,760 clones, was constructed using ZAP Express cDNA Synthesis Kit (STRATAGENE). First, each insert in the pBK-CMV vector was sequenced by the 5' sequencing method using the T3 primer. The sequencing results of cDNAs from the garlic library were part of the data used for establishing the publically available on-line garlic EST database (Kim et al. 2009). By carrying out a BLAST search against the Uni-prot database with the cDNA sequences from our garlic cDNA library, we identified candidate clones for the cellulase, endo-glucanase, exo-glucanase, and glucosidase genes. Among them we have selected a cDNA clone, named ASS8P20, which has a full-length cDNA sequence of a novel garlic endo-beta-1,4-glucanase gene. The nucleotide sequence data of the novel garlic endo-beta-1,4-glucanase gene reported in this paper has been deposited in the DDBJ/EMBL/GenBank databases under accession number AK343185.

### Cloning of a Garlic Cellulase Gene

The open reading frame (ORF) of the ASS8P20 clone was predicted using the NCBI ORF finder. The ORF region of ASS8P20 was amplified using two designed primers, 8P20-

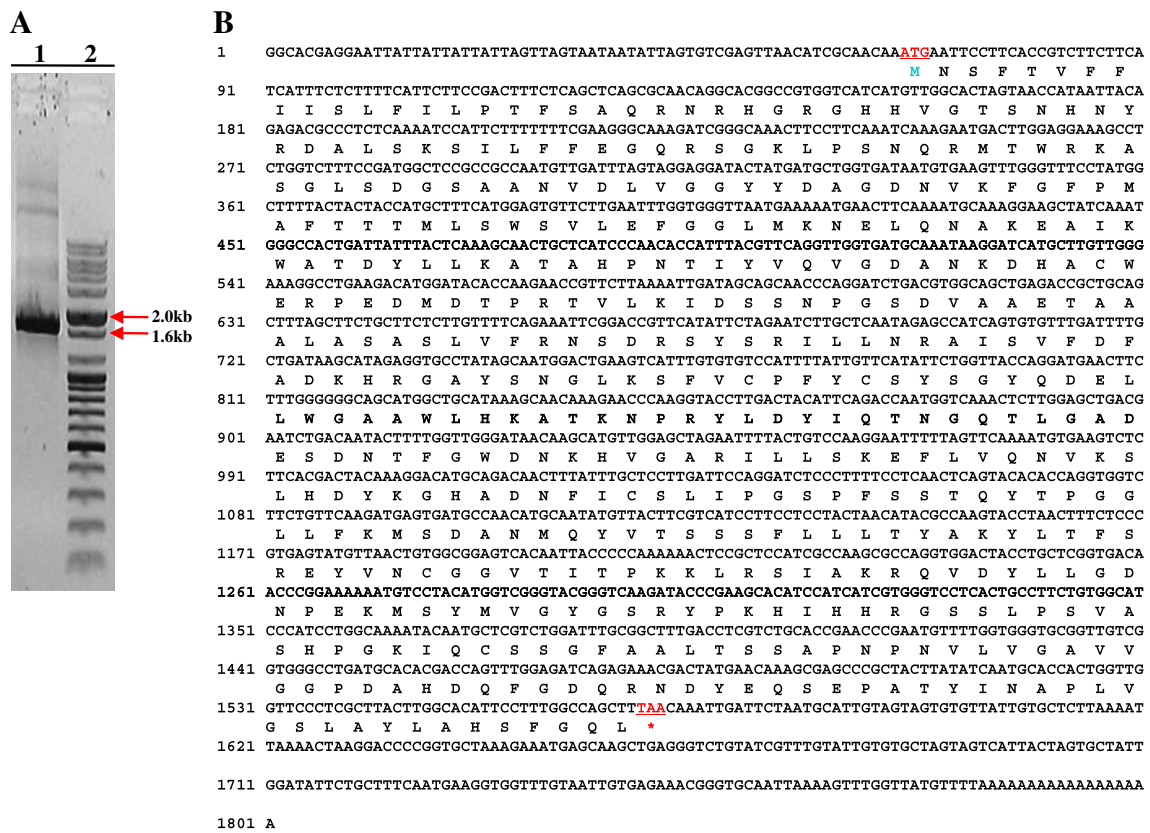
NdeI (5'-GCCATATGAATTCCTTCACC-3') and 8P20-BamHI (5'-AAGGATCCTTAAAGCTGGCCAAAGGAAT-3'). These two primers were designed to include specific enzyme sites, NdeI (for 8P20-NdeI) and BamHI (for 8P20-BamHI), respectively, for sub-cloning into the expression cloning vector, pET28a. The amplification of the ORF region was performed using an ABI-polymerase chain reaction (PCR) machine. The 20- $\mu$ l reaction volume contained 0.5  $\mu$ l Taq DNA polymerase (SolGent), 2  $\mu$ l 10 $\times$  Taq buffer, 2.5 mM dNTP mix, 3.5 pmol of each primer, and 1  $\mu$ l cDNA template. PCR amplification was carried out for 32 cycles, after the initial hot start incubation at 96°C (5 min). Each cycle included a denaturing step at 95°C (30 s), an annealing step at 55°C (45 s) and an elongation step at 72°C (4 min). The terminal incubation was at 72°C (5 min). The resulting PCR product was ligated into the T&A cloning vector (RBC Bioscience) by incubation at 4°C overnight in a 10- $\mu$ l reaction volume. The *Escherichia coli* DH5 $\alpha$  strain was transformed with 5  $\mu$ l of the ligation product. Plasmid harboring the 1,506-bp insert was sequenced using two primers, M13 and M13 reverse primers, present on the RBC T&A cloning vector.

### Sub-cloning of the Cellulase Gene Into the pET28a Expression Vector

The sequenced ORF of ASS8P20 was removed from the RBC T&A cloning vector by digestion with the restriction enzymes BamHI and NdeI. It was subcloned into pET28a, the MCS of which includes BamHI and NdeI restriction sites. The recombinant pET28a plasmid was transferred into *E. coli* DH5 $\alpha$ . Among the candidate transformants, 10 colonies were chosen and cultured. The recombinant plasmids were isolated and digested with BamHI and NdeI to confirm the insertion of the ORF sequence.

### Expression Assay of Total Proteins

Recombinant pET28a plasmids with the correct inserts were transferred into *E. coli* BL21. The transformed *E. coli* was cultured in a 50-ml conical tube in 5 ml of LB (Km<sup>+</sup>) until an OD<sub>600</sub> of 0.5 was reached. Then 5  $\mu$ l IPTG (1 mmol/ $\mu$ l) was added to the culture, and the cultivation was continued at 37°C, 230 rpm for 3 more hours. The control *E. coli* BL21 strain, harboring an empty pET28a expression vector, was cultured in the same way for comparison. To identify the expression of total proteins, 20  $\mu$ l of culture was centrifuged at 4°C, 8,000 rpm for 6 min. The pellet was resuspended in 1 M Tris-HCl (pH 6.8). To check the solubility of the protein, 200  $\mu$ l of culture was pelleted at 4°C, 8,000 rpm for 6 min and the pellet was resuspended in 1 M Tris-HCl (pH 6.8). The cell walls of the resuspended cells were sonicated until the



**Fig. 1** Nucleotide sequence of ASS8P20 cDNA and its deduced amino acid sequence. **a** Agarose gel electrophoresis of the ASS8P20 full-length cDNA fragment (EcoRI/XhoI) from the garlic cDNA library shows a band of about 1.8 kb. **1** 1.8 kb full-length cDNA of

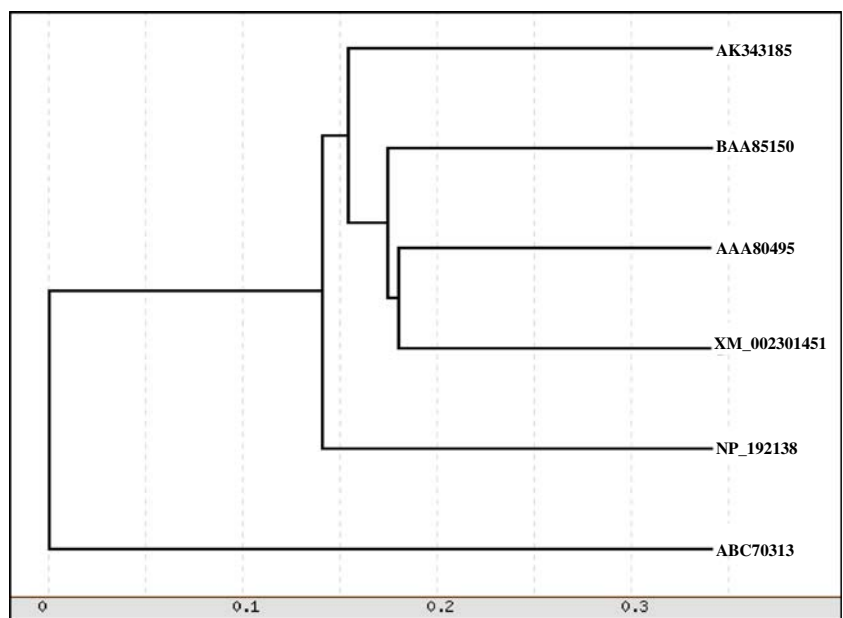
ASS8P20, **2** DNA size marker. **b** The full-length cDNA sequence of ASS8P20 comprises 1,801 bp and an ORF coding region, which includes 501 deduced amino acids. \* stop codon. The start and stop codons of the ORF are *underlined*

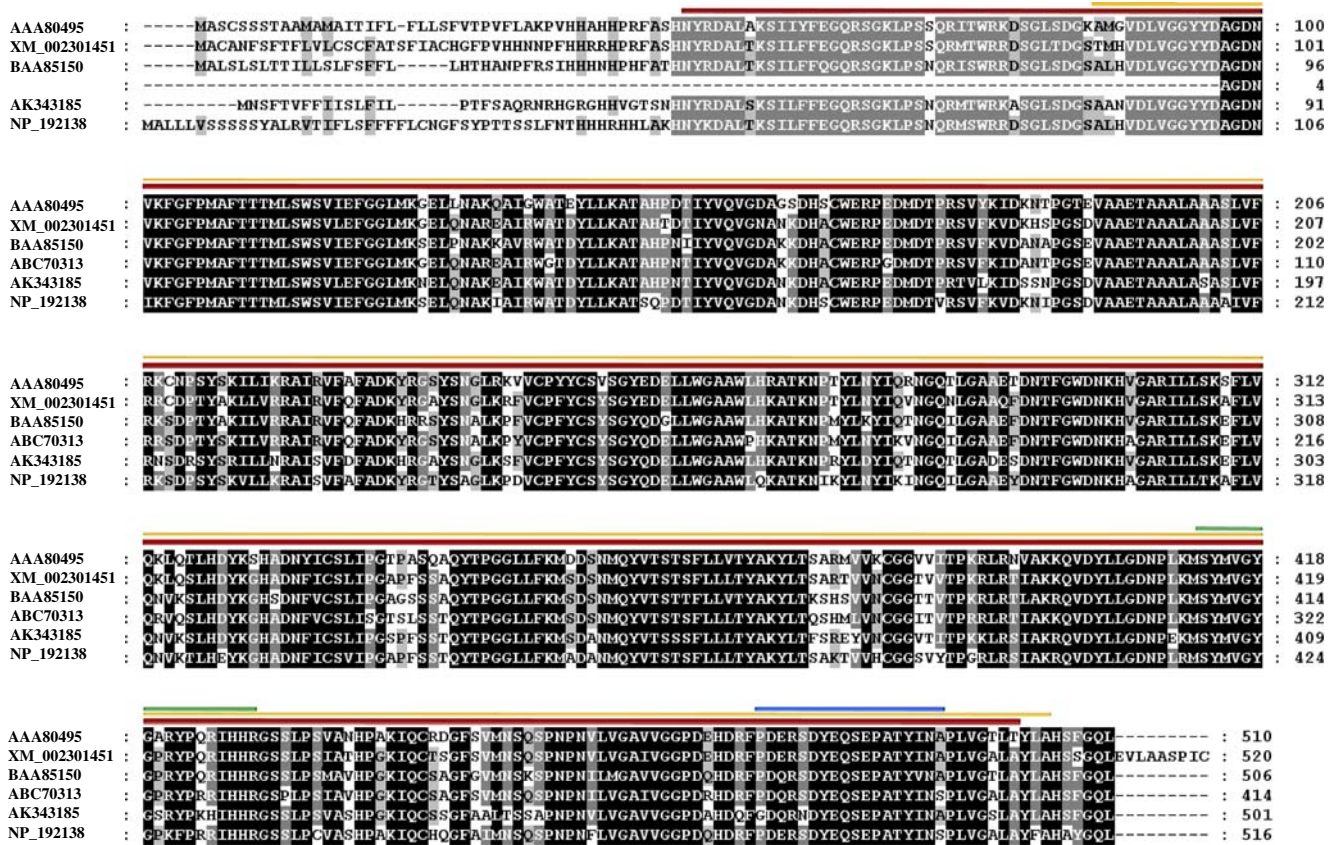
suspension became clear. Then the suspension was centrifuged at 4°C, 10,000 rpm for 20 min, and the supernatant was removed to another clear Eppendorf tube. The pellet was resuspended in 150 µl of 1 M Tris–HCl (pH 6.8).

Protein Expression Assay with SDS-PAGE

The molecular weight of the cellulase protein was determined by sodium-dodecyl sulfate–polyacrylamide gel electrophoresis

**Fig. 2** Phylogenetic tree using amino acid sequences of endo-beta-1,4-glucanase homologs from garlic (AK343185), *P. sativum* (BAA85150), *G. max* (ABC70313), *A. thaliana* (NP\_192138), *P. trichocarpa* (XM\_002301451), and tomato (AAA80495)





**Fig. 3** Multiple alignment indicates that ASS8P20 is a novel garlic endo-beta-1,4-glucanase gene. The multiple alignment has been carried out among amino acid sequences of endo-1,4-beta-glucanase homologs from garlic (AK343185), *P. sativum* (BAA85150), *G. max* (ABC70313), *A. thaliana* (NP\_192138), *P. trichocarpa* (XM\_002301451),

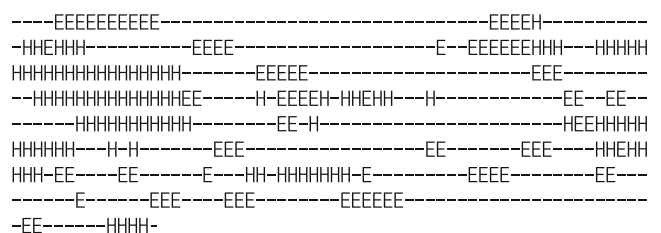
and tomato (AAA80495). *Red-brown bar* glycoside hydrolase family 9, *orange bar* endo-beta-1,4-glucanase subfamily, *green bar* active site 1 of glycoside hydrolase family 9, *blue bar* active site 2 of glycoside hydrolase family 9

(SDS-PAGE) at room temperature, using NEXT GEL™ acrylamide resolving gel with 1× NEXT GEL™ RUNNING BUFFER. Protein size and percentage was determined using Precision Plus Protein Dual Color Standards (Bio-Rad). After protein separation, the gel was stained with Coomassie blue staining solution containing 45% methanol, 45% H<sub>2</sub>O, and 10% acetic acid and brilliant blue R (Sigma-Aldrich) for 20 min. The gel was destained using destaining buffer (45% methanol, 45% H<sub>2</sub>O, and 10% acetic acid) until the protein bands appeared clearly on the gel.

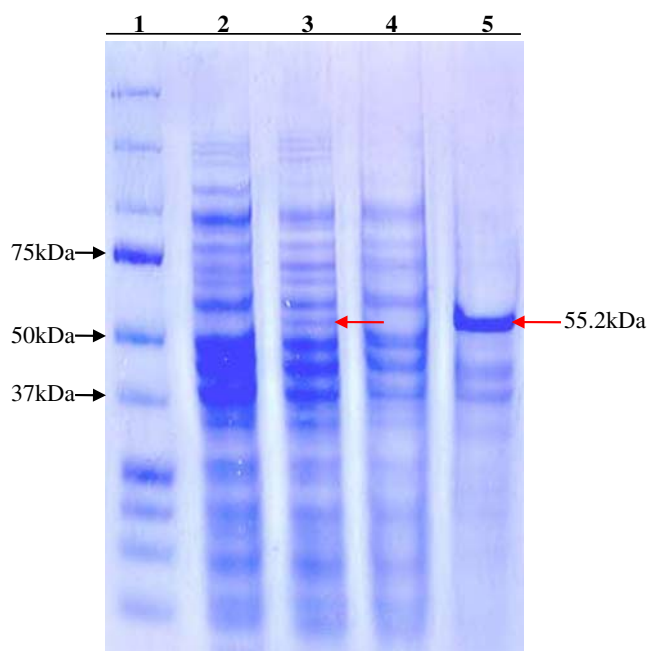
**Assay of Cellulase Activity**

Cellulase protein activity was measured using EnzChek cellulase substrate (Invitrogen). The cellulase substrate stock solution was prepared by adding 1.86 ml of 50% DMSO. To make the working solution, the cellulase substrate stock solution was diluted fivefold by adding 7.44 ml of digestion buffer (100 mM sodium acetate, pH 5.0). For preparing the enzyme standard and sample dilutions, four dilutions of standard cellulase isolated from *Trichoderma reesei* ATCC 26921 (SIGMA Catalog No.

C2730) were prepared in 50 µl digestion buffer. Dilutions of the standard contained 0.6, 0.8, 1.0, or 1.2 µU/µl of cellulase. Dilutions of the test samples were also prepared in 50 µl of digestion buffer and contained 5, 10, 20, or 50 µl of the test sample, respectively. Then 50 µl of the working substrate solution and 50 µl of each cellulase dilution were mixed and incubated for 30 min at room temperature. Cellulase activity was assayed by measuring fluorescence using a microplate reader. Excitation and emission maxima were observed at 339 nm (shoulder at 360 nm) and 452 nm.



**Fig. 4** Secondary structure of ASS8P20. We constructed the secondary structure of ASS8P20 using NNPREDPREDICT (<http://alexander.compbio.ucsf.edu/~nomi/nnpredict.html>). H helix, E strand, - no prediction



**Fig. 5** Analysis of the protein coded by the ASS8P20 ORF and purification of the recombinant protein. The recombinant protein was identified by SDS-PAGE. *Lane 1* protein molecular weight marker, *lane 2* total protein of *E. coli* BL21(DE3) pLysS harboring empty pET-28a as a control, *lane 3* total protein in the *E. coli* cells harboring the recombinant vector when induced by 0.5 mM IPTG, *lane 4* soluble proteins, *lane 5* insoluble proteins. Red arrows indicate the recombinant protein with a molecular weight of 55.2 kDa

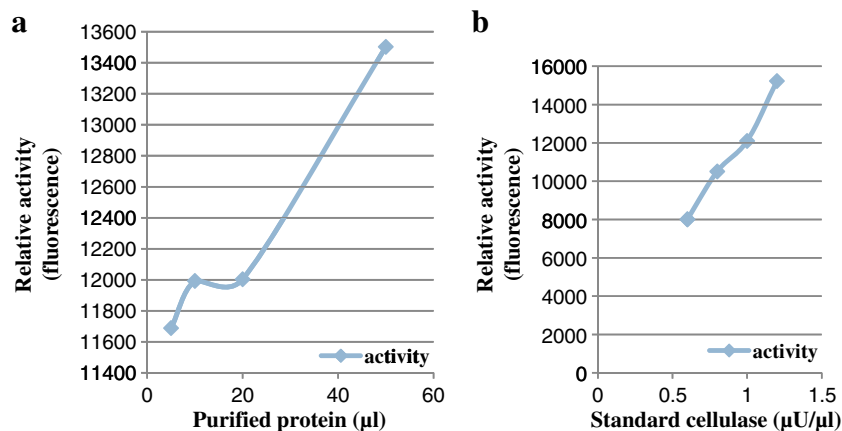
## Results and Discussion

### Identification of a Novel Garlic Endo-Beta-1,4-Glucanase Gene

In this study we constructed a cDNA library using total RNA isolated from garlic and sequenced 5,760 cDNA

clones from it. Some of our sequencing results contributed to the establishment of the garlic EST database, which was published previously (Kim et al. 2009). Among our garlic cDNA clones, we identified one clone, named ASS8P20, with a cDNA sequence having a high similarity and identity with those of the previously known endo-beta-1,4-glucanase homologs in other organisms. Using agarose gel electrophoresis, we confirmed that the full-length cDNA fragment, which had been cut from the pBK-CMV vector with the restriction enzymes EcoRI and XhoI, was about 1.8 kb. We also found that the ASS8P20 full-length cDNA sequence is 1,801 bp in nucleotide length (Fig. 1). Its ORF contained 1,506 bp, corresponding to 501 amino acids with a predicted molecular weight of 55.2 kDa. The deduced amino acid sequence of ASS8P20 exhibited 81% identity with that of the *Pisum sativum* endo-beta-1,4-glucanase (BAA85150), 81% with that of *Glycine max* endo-beta-1,4-glucanase precursor (ABC70313), 77% with that of the *Arabidopsis thaliana* glycosyl hydrolase 9B13 (NP\_192138), and 75% with that of the tomato endo-1,4-beta-glucanase precursor (AAA80495). According to the analysis of the phylogenetic tree, the garlic ASS8P20 is very close to the *P. sativum* endo-beta-1,4-glucanase (BAA85150; Fig. 2).

In addition, we carried out multiple sequence alignments of these amino acid sequences and searched for domains and motifs that would be crucial for the molecular functions of the garlic ASS8P20. Surprisingly, the deduced amino acid sequence of our garlic ASS8P20 gene perfectly matched domains of the glycoside hydrolase family 9 and endo-beta-1,4-glucanase subfamily, including active site 1 and active site 2 of glycoside hydrolase family 9, which are very common and typical of endo-beta-1,4-glucanase enzymes, compared with other sequences (Fig. 3). This demonstrates that our garlic ASS8P20 gene must be



**Fig. 6** Activity assay of novel garlic endo-beta-1,4-glucanase protein. The reaction mixture included 50 µl enzyme solution (0.02, 0.04, 0.08, and 0.2 µg/µl in four test enzyme solutions in terms of dry protein weight, respectively) and 50 µl substrate solution (0.2 mM carboxymethyl cellulose in 100 mM sodium acetate, pH 4.6). The

reactions were incubated for 30 min at room temperature. **a** The volumes of purified protein stock solution (0.2 µg/µl) was 5, 10, 20, and 50 µl in the reaction mixtures. **b** The standard cellulase activity was measured using 0.6, 0.8, 1.0, and 1.2 µU/µl

encoding a novel endo-beta-1,4-glucanase. Also, using the website (<http://alexander.compbio.ucsf.edu/~nomi/npredict.html>), we predicted a secondary structure for the garlic ASS8P20 amino acid sequence and found that there are several parts of helical strands in this protein (Fig. 4), implying that this novel protein would be involved in biological functions in a mechanically stable form.

#### In Vitro Analysis of Expression of This Garlic ASS8P20 Gene Into Protein

We designed specific primers that included BamHI and NdeI restriction enzyme sites for cloning the ASS8P20 ORF into the pET28a vector. Then we ligated the PCR products into a pGEM-T *easy* TA cloning vector and transformed *E. coli* DH5 $\alpha$  with the recombinant vector. After the correct recombinant plasmids were identified by digestion with the restriction enzymes BamHI and NdeI and sequencing the selected inserts, the plasmids were transferred into *E. coli* BL21 to assess whether the gene encodes the garlic cellulase protein. SDS-PAGE analysis of total proteins from the transformed *E. coli* BL21 showed a single strong band of a novel protein in an insoluble form, which was about the predicted molecular weight of 55.2 kDa. In contrast, this band was absent from the total proteins from *E. coli* BL21 harboring the empty vector pET28a (Fig. 5). The protein was purified using a Ni-NTA His-Tag protein purification column and was used in the next step.

#### Cellulase Activity Assay of the Garlic ASS8P20 Protein

The cellulase activity of the garlic ASS8P20 protein was tested, as described in the “Materials and Methods”. The reaction mixture included 50  $\mu$ l enzyme solution (0.02, 0.04, 0.08, and 0.2  $\mu$ g/ $\mu$ l in four test enzyme solutions in terms of dry protein weight, respectively) and 50  $\mu$ l substrate solution (0.2 mM carboxymethyl cellulose in 100 mM sodium acetate, pH 4.6). The reaction was incubated for 30 min at room temperature. The volumes of purified protein stock solution (0.2  $\mu$ g/ $\mu$ l) were 5, 10, 20, and 50  $\mu$ l in the reaction mixtures (Fig. 6a). The activity of the standard cellulase was measured at concentrations of 0.6, 0.8, 1.0, and 1.2  $\mu$ U/ $\mu$ l (Fig. 6b). The ASS8P20 protein exhibited strong cellulase activities, which were approximately equivalent to those of 50  $\mu$ l enzyme solutions at 0.8, 1.0, and 1.2  $\mu$ U/ $\mu$ l of the standard cellulase (Fig. 6). This result clearly demonstrates that the garlic ASS8P20 gene encodes a novel cellulase protein.

In this study, we have identified a novel garlic cellulase gene using our garlic cDNA library. Our analysis of the protein sequence showed a high degree of identity with the

sequences of endo-beta-1,4-glucanase homologs from *P. sativum*, *Solanum lycopersicum*, *Populus trichocarpa*, *A. thaliana*, and *G. max*. The sequences exhibited perfectly matched regions with the functional domains of the previously known endo-beta-1,4-glucanases. In addition, our in vitro protein expression assay of the garlic ASS8P20 gene showed that it was translated into a novel protein of the predicted molecular weight of 55.2 kDa. Furthermore, the assay of the cellulase activity of the protein clearly showed that it is a cellulase protein. In fact, although several investigators previously reported that under natural conditions, the cell walls of the garlic could be softened and disorganized and that extracts obtained from garlic had cellulase activity, little was known about the actual existence of the garlic cellulase gene. Therefore, our study is the first experimental evidence for the garlic cellulase at the molecular level and further reinforces the usefulness of garlic not only to the pharmaceutical and culinary industries, but also as a source of important genes including the cellulase gene, which would be applicable to several industrial processes.

**Acknowledgements** This study was supported by grant 2007-04269 from the Ministry of Education, Science in Korea.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

#### References

- Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11:113–116
- Del CE, Bennett AB (1996) Pedicel breakstrength and cellulase gene expression during tomato flower abscission. *Plant Physiol* 111(3):813–820. doi:111/3/813 [pii]
- Ferrarese L, Trainotti L, Moretto P, Polverino DLP, Rascio N, Casadoro G (1995) Differential ethylene-inducible expression of cellulase in pepper plants. *Plant Mol Biol* 29(4):735–747
- Juan F, Antonia H, Concepcion S, Rocio R, Rafael G, Ana J (2001) Effect of dressings “(alinos)” on olive texture: cellulase, polygalacturonase and glycosidase activities of garlic and lemon present in brines. *Eur Food Res Technol* 212:465–468
- Kim DW, Jung TS, Nam SH, Kwon HR, Kim A, Chae SH, Choi SH, Kim RN, Park HS (2009) GarlicESTdb: an online database and mining tool for garlic EST sequences. *BMC Plant Biol* 9:61. doi:1471-2229-9-61 [pii] doi:10.1186/1471-2229-9-61
- Krause DO, Denman SE, Mackie RI, Morrison M, Rae AL, Attwood GT, McSweeney CS (2003) Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics. *FEMS Microbiol Rev* 27:663–693. doi:S016864450300072X [pii]
- Osborne DJ, Jackson MB, North Atlantic Treaty Organization. Scientific Affairs Division (1989) Cell separation in plants: physiology, biochemistry, and molecular biology. Springer, Berlin