Cloning and characterization of a novel trypsin inhibitor (BTIw1) gene from Fagopyrum esculentum

YUYING LI, ZHENG ZHANG, AIHUA LIANG, & ZHUANHUA WANG

Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Institute of Biotechnology, Shanxi University, Taiyuan 030006, People’s Republic of China

(Received 14 December 2005)

Abstract
Based on the amino acid information of trypsin inhibitor of buckwheat (Fagopyrum Esculentum Moench), degenerated primers were designed and a full-length cDNA sequence named BTIw1 (Buckwheat Trypsin Inhibitor) was amplified from the leaves RNA by using RT-PCR and rapid amplification of cDNA ends (RACE) methods. Sequence analysis shows that the 392 bp cDNA contained an open reading frame (ORF) of 216 bp, encoding 72 amino acids residues. The deduced amino acid sequence exhibits 96 and 93% homology with BWI-1 and BTI-2, a natural trypsin inhibitor from buckwheat seeds. Southern blotting suggested that three copies of BTIw1 gene existed in the buckwheat genome. Moreover, a predicted secondary structure and 3D-structural model was constructed by homology modeling. To our knowledge, this is the first all-round report of the gene BTIw1. The novel BTIw1 gene has been submitted to the GeneBank under Accession No. DQ289792.

Keywords: BTIw1 gene, buckwheat, cDNA sequence, RACE, proteinase inhibitor

Introduction
Protease inhibitor is widely distributed in nature including many kinds of animals, plants and microorganisms. It plays key regulatory roles in many biological processes, including the blood coagulation system, the complement cascade, apoptosis and the hormone processing pathways (Neurath 1989). Most inhibitors were found to regulate the proteolytic activity of the target enzymes. Plant protease inhibitor is present in almost every kind of plant, and mainly concentrated in plant seeds and tubers. They fight against insects and fungi attacks by suppressing the activities of exogenous proteases of pathogens and pests. It has been identified as a defense protein developed in plants as well as a possible anti-nutritional factor (Bowles 1990).

Buckwheat is widely cultivated in Asia, Eastern Europe, North America and Australia. Buckwheat seeds contain many valuable substances with high levels of nutritional components. Many studies have been reported on proteins, carbohydrates, vitamins, fibers, enzymes and other bio-active substances (Pomeranz 1983; Skrabanja et al. 2001). Buckwheat has many advantages in nutrition and medicine. Its leaf, stem, flower and seeds all contain healthy components which are used to treat choking, ulcer, haemostatic and even hypertension and diabetes, according to the Chinese Material Medicine Dictionary. Protease inhibitor, a general kind of protective mechanism during seed germination and plant development, also weighed a lot in buckwheat seeds. Buckwheat can obtain self-regard protection against insect predators and against environment hazards during germination and seedling growth. For the whole grown plant, inhibitors can be secreted induced by the injured parts, and thus to fight against invasion of pests.

Although a number of trypsin inhibitors have been characterized and their amino acid of some trypsin
inhibitors have been sequenced from buckwheat seeds (Ikeeda and Kusano 1983; Belozersky et al. 1990), on the level of gene for biomolecules in buckwheat, relatively little research has been done. Till now, the nucleotide sequences have not been determined. In this paper we sequenced and characterized a full-length cDNA of trypsin inhibitor from buckwheat leaves, also we compared its deduced amino acid sequence with those from different representative protease inhibitor of plant. This work is to provide new and significant information in the scientific research of buckwheat, and it will be useful for further study of buckwheat in the fields of nutrition, medicine and agriculture exploration and development and so on.

Materials and methods

Plant materials and reagents

Common Buckwheat (*Fagopyrum Aesculentum Moench*) seeds were obtained from Shouyang of Shanxi province (China). Seeds were germinated ten days, and the first leaves were collected and immediately used for DNA and RNA isolation. 3'-RACE kit was purchased from Invitrogen (USA); pGEM®-T Easy Vector was purchased from Promega (USA)

**RNA isolation and cDNA synthesis**

The total RNA was isolated from buckwheat leaves, according to the method described previously by Li et al. (2004). The concentration was determined by the absorption at 260 and 280 nm. Total RNA was then used in RT-PCR and RACE to amplify the sequence of 3' and 5' end of the cDNA. First strand cDNA (cDNA I) of *BTIw1* was synthesized using the isolated total RNA and Oligo dT primer for reverse transcription. Nucleotide end sequence of 3'-end of the *BTIw1* cDNA (cDNA II) were amplified by the method of 3' RACE. The primers used for 3' RACE were Oligo (dT)_{17} adaptor primer AP: 5'-GGC CAC GCG TCG ACT AGT ACT TTT TTT TTT TTT T-3' (for first strand cDNA synthesis). Amplification of 5'-end was performed by using 5'-RACE Kit, and gene specific primer GP_{4} (5'-AAA CCC ACA CAC GGT CAC AT-3'), which was synthesized for first strand cDNA (cDNA III).

**RTPCR cloning partial *BTIw1* gene**

According to the amino acid sequences of reported trypsin inhibitors from common buckwheat (Belozersky et al. 1995, Dunaevsky et al. 1997), the degenerate sense primer and antisense primer L1: 5'-GG(TA)AAACAAAGA(AG)-TGGCC(TA)GA(AG)CT-3' and L2: 5'-(AG)AT(AG-T)GC(AG)CG(AC)ACATC(TC)TC(GA)TTTTC-3' were designed. The primers were used to amplify partial *BTIw1* sequence using cDNA I as a template. PCR was performed on a Peltier thermal gradient cycler-200 using a program of pre-incubation at 94°C for 4 min; 5 cycles consisting of denaturation at 94°C for 1 min, annealing at 42°C for 1 min, and extension at 72°C for 1 min; followed by 30 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min, and finally post-treatment for 10 min at 72°C. The PCR products were electrophoresced on 1.5% agarose gel.

**Rapid amplification of cDNA ends (RACE)**

3'end sequence of *BTIw1* was amplified using 3'RACE kit following the user manual. At first, amplification was performed in a 50 μl volume using cDNA II as a template, the degenerate primer L1 and abridged universal amplification primer AUAP (5'-GGC CAC GCG TCG ACT AGT AC-3') as primer. Reaction conditions were 94°C for 4 min, followed by 5 cycles of 94°C for 1 min, 42°C for 1 min and 72°C for 1 min, and then, followed by 30 cycles of 94°C for 1 min, 56°C for 1 min and 72°C for 1 min, and finally post-treatment for 10 min at 72°C. Nested PCR amplification was performed with AUAP and gene-specific primer GP1 (5'-GAGAGAGGTTCAAGGCTG-CCAAGAT-3'), using the product of the first PCR amplification as template. Reaction conditions were 94°C for 4 min and followed by 30 cycles of 94°C for 1 min, 60°C for 1 min and 72°C for 1 min, and finally post-treatment for 10 min at 72°C.

5'end sequence of *BTIw1* was amplified using 5'RACE kit. Using cDNA III as template, upstream adaptor amplification primer AAP (5'-GGCCACGC-GTCGACTAGTACGTTTTTTTGTTCGTTCTCGTTTT-3') and gene specific primer L6 (5'-GGCTACCTCGAGCAAG-3') as primers. PCR was performed in 50 μl volume under the following condition: template was denatured at 94°C for 4 min and followed by 4 cycles of 94°C for 1 min, 42°C for 1 min and 72°C for 1 min, followed by 30 cycles of 94°C for 1 min, 60°C for 1 min and 72°C for 1 min, and finally post-treatment for 10 min at 72°C. And then, using the product of the first PCR as template, gene specific primer L5 (5'-GGCTACCTCGAGCAAG-3') and AUAP as primers, Nest PCR was performed under the condition: template was denatured at 94°C for 4 min and followed by 30 cycles of 94°C for 1 min, 62°C for 1 min and 72°C for 1 min, and finally post-treatment for 10 min at 72°C.

**Cloning of the full-length cDNA of *BTIw1***

A full-length clone of *BTIw1* was created by joining the partial *BTIw1*, 5', and 3' fragments using overlapping PCR. Several fragments of cDNA of *BTIw1* were purified according to the handbook of Gel Extraction Mini Kit manufactured by Watson Biotechnologies, Inc. All PCR fragments were sub-cloned into pGEM®-T Easy Vector and were transformed into *E. coli* DH5a. Recombinant plasmid DNA was isolated according to Watson's handbook of Plasmid Mini Kit recommended.
by the manufacturer. Nucleotide sequences were determined by an automatic DNA sequencer (ABI prism, USA).

Genomic DNA extraction and Southern blot analysis

Genomic DNA was isolated from buckwheat leaves and stems using an improved procedure based on methods described by Dellaporta et al. (1983). The DNA concentration was determined by the absorption at 260 and 280 nm. The genomic DNA was first digested with 5 U of EcoRV. After electrophoresis on 0.8% agarose gels, DNA was blotted on a Pyroxylin membrane in 20 × SSC buffer (3.0 M NaCl; 0.3 M sodium citrate, pH 7.0) and 10% SDS for 20 min. DNA was cross-linked to the membrane using CL-1000 Ultraviolet Cross-linker (UVP). Pre-hybridization and hybridization was performed in Pre-hybridization buffer (0.25 M Na2HPO4, 1 mM EDTA, 7% SDS, pH 7.2, 0.5% blocking reagent) at 68°C for 6 h, the membrane was hybridized with denatured cDNA (216 bp-length fragment of BTIw1 gene) as a probe at 68°C overnight. Moreover, the probe was labeled with digoxigenin DIG DNA labeling system (Roche, Basel, Switzerland). Then the membrane was incubated with the dye-NBT/BCIP for 15 min.

Results and discussions

Synthesis and sequence of BTIw1 cDNA

Three fragments of cDNA (90, 350 and 200 bp) of BTIw1 gene were amplified from the leaves of 10-day-old. The PCR strategy is shown in Figure 1. From overlapping sequences of our fragments the full-length cDNA BTIw1 was constructed. The nucleotide sequence and the deduced amino acid sequence are shown in Figure 2. Sequence analysis indicated that the full-length cDNA contains an open reading frame of 392 bp encoding for 72 amino acids residues. A start codon (ATG) was at nucleotide position 30–32, a stop codon (TGA) was at 245–248. Two potential polyadenylation signals (AATAAA) were found at positions 300 and 361. A start codon sequence AAGATGGG was found accorded with Kozak conserved character (it is A/G at position −3, it is G at position +4), promoter elements in plant genes were presently poorly defined. Here we have not found a CAAT, AGGA or TATA box sequence (Pelham 1982) in BTIw1. By being subjected to the BLAST search program (http://www.ncbi.nlm.nih.gov/BLAST/), the obtained nucleotide sequence from leaves of common buckwheat has no homology with other plants’ protease inhibitor completely. Furthermore, the deduced amino acid sequence of BTIw1 was compared with the amino acid sequences from several other plants, using BOX-SHADE 3.21 server (http://www.ch.embnet.org/software/BOX_form.html). It shares 96 and 93% homology with natural BWI-1 and BTI-2 trypsin inhibitor from Fagopyrum esculentum, 67% homology with ATI trypsin inhibitor from Amaranthus hypochondriacus, 58% homology with Luti trypsin inhibitor from Linum Usitatissimum respectively (see Figure 3). We named the nucleotide sequence as BTIw1, the novel BTIw1 gene has been submitted to the GenBank under Accession No. DQ289792.

The result of Southern blotting showed three intensive bands in both leaves and stems of buckwheat (see Figure 4). It suggests that buckwheat trypsin inhibitor genome might contain at least three copies of the BTIw1 gene.

Computer-assisted analysis

The secondary structure of BTIw1 was predicted with the PSIPRED Protein Structure Prediction Server (http://bioinf.cs.ucl.ac.uk/psipred) and the results were collected and analyzed by comparing the data manually. BTIw1 consisted of 1 α-helix and 5 β-sheets (extended) connected with coils (loop). Secondary structure of BTIw1 protein showed 13.89% α-helix, 27.78% β-sheet and 58.33% blank (other loop). The BTIw1 protein was identified from the residues, it has 76.39% residues exposed in molecular outside surface, 23.61% residues distributed inside. All of these suggested that BTIw1 protein might be a water-soluble protein.
The three-dimensional structure was analyzed with Predict Protein structure prediction and sequence analysis server (http://cubic.bioc.columbia.edu/pp/). BTIw1 protein is indicated apparently not as globin according to globin Predict. These results are very useful for studying novel protein on structure and function.

Phylogenetic tree was drawn using Clustalx 1.83 and displayed with Njplot based on the homology of the amino acid sequences of various proteinase inhibitor types. Results of phylogenetic analysis revealed diversification into Monocotyledon and Dicotyledon groups (Figure 5).

The model predicts only one pair of disulfide bonds from between Cys7 and Cys52. BTIw1 protein sequence show high contents of Val (16.7%), Glu (11.1%), and Arg (9.7%). These are in good agreement with the amino acid composition of BWI-1 isolated by Beloizersky et al. (1995) and BTI-2 isolated by Pandya et al. (1996). However, at N-terminus of BTIw1 three new amino acids including Met, Ala and Try were found. Furthermore, Gly45 in BTIw1 is replaced by Ala42 in BWI-1 or BTI-2. Gly45 was highly conserved in other serine proteases inhibitor. Gly and Ala are nonpolar aliphatic amino acids, so the difference has little influence on the character of BTIw1 protein. Using Signal P 3.0 Server for signal peptide predication, signal peptide was not found. In addition, the buckwheat trypsin inhibitor shows strong sequence similarities with the potato I family of serine protease inhibitors.

![Figure 2](image)

**Figure 2.** Nucleotide sequence and its deduced amino acid sequence of a cDNA encoding for BTIw1. The polyadenylation signals are single line underlined (—) and the stop codon is indicated by an asterisk (*)

![Figure 3](image)

**Figure 3.** Alignment of amino acid sequences deduced from nucleotide sequences of BTIw1 and other plant proteinase inhibitor reported in existing gene databases on internet with the computer program BOXSHADE (ver. 3.21). The identical and similar residues are shown in reverse type and shaded boxes, respectively. The text in the figure refers to the entry name of Uniprot Proteins databanks. The arrow indicates the putative reactive (inhibitor) site residues at Arg48 and Asp49.
Cloning and characterization of BTIw1

Richardson 1991). It contains a conserved potato inhibitor I family signature WPELVGERGSKA.

Conclusions

We have cloned a full-length cDNA of BTIw1 gene. The amino acid sequence the BTIw1 protein is highly conserved when compared with other species protease inhibitor. It can be considered a member of potato I family of serine protease inhibitors. This is the first notification of the nucleotide sequence representing buckwheat trypsin inhibitor. This work will be useful for further study on the molecular mechanism of buckwheat trypsin inhibitor in the fields of nutrition, medicine and plant self-protection and so on. Further studies are being carried out to examine the bioactivity of BTIw1.

Acknowledgements

This study was supported by the Natural Sciences Foundation of China (grant No. 30470178) and the Natural Sciences Foundation of Shan Xi Province (grant No. 20031064).

References