

**Plant Gene Register**

# Nucleotide and Deduced Amino Acid Sequence of the 22-Kilodalton Cathepsin D Inhibitor Protein of Potato (*Solanum tuberosum* L.)<sup>1</sup>

David J. Hannapel\*

Department of Horticulture, Iowa State University, Ames, Iowa 50011

Several storage proteins of potato (*Solanum tuberosum* L.) tubers have been identified, including several proteinase inhibitors such as proteinase inhibitor I and II and the 22-kD protein group (Melville and Ryan, 1972; Bryant et al., 1976; Suh et al., 1990). The 22-kD cathepsin D inhibitor protein of potatoes (PDI) has been purified and characterized (Mares et al., 1989). This glycosylated protein inhibits cathepsin D and trypsin activity and is homologous with the soybean trypsin inhibitor family (Mares et al., 1989). Recent studies showed that PDI transcripts accumulate in potato tubers, stems, roots,

and leaves depending on genotype and stage of development (Hannapel, 1991). Accumulation of potato cathepsin D inhibitor mRNA is wound inducible in leaves and occurs both locally and systemically (Suh et al., 1991). The PDI cDNA clone p749 was isolated from a tuber cDNA library using differential screening (Table I). The DNA sequence data and the deduced amino acid sequence are shown in Figure 1. The triangle indicates the site of cleavage for the signal peptide, and the first asterisk at Arg<sup>99</sup> marks the trypsin-binding reactive site. The second asterisk marks the TAG stop codon, and the polyadenylation signals are underlined. The deduced amino acid sequence has 92 and 87% homology, respectively, to a purified PDI (Mares et al., 1989) and another novel inhibitor of cathepsin D (Ritonja et al., 1990). The nucleotide sequence of p749 has 94% homology to an aspartic proteinase inhibitor homolog from potato tubers (Strukelj et al., 1990).

<sup>1</sup> This work was supported by the Iowa State University Biotechnology Council. Journal Paper No. J-14950 of the Iowa Agriculture and Home Economics Experiment Station, Ames, IA. Project No. 3056.

\* Fax 1-515-294-0730.

**Table 1.** Characteristics of a cDNA encoding the cathepsin D inhibitor protein of potato

**Organism:**  
Potato (*Solanum tuberosum* L. cv Superior).

**Location in Genome:**  
Nuclear genome.

**Gene, Function, Pathway:**  
p749; inhibitor of aspartic proteinases.

**Techniques:**  
cDNA library screening; double-stranded plasmid sequencing of both strands; subcloning of cDNA fragments.

**Method of Identification:**  
Amino acid sequence comparison with a cathepsin D inhibitor from potato.

**Expression Characteristics:**  
Transcript of about 1100 nucleotides. Transcripts accumulate in leaves in response to wounding and high levels of sucrose; transcripts abundant in potato tubers.

**Features of cDNA structure:**  
Contains an open reading frame of 663 bp.

**(G + C) Content:**  
37.7%.

**Structural Features of the Protein:**  
An open reading frame of 221 amino acids; *M<sub>r</sub>*, 28,430; signal peptide cleavage site at Glu<sup>33</sup>; trypsin-binding reactive site at Arg<sup>99</sup>; protein is most abundant in tubers but accumulates in leaves in response to wounding.

1	ATG ATG AAG TGT TTA TTT TTG TTA TGT CTG TGT TTG TTT CCC ATT TTG GTA TTT TCA TCA	20
	M M K C L F L L C L C L F F I L V F S S	
61	ACT TTC ACT TCG CAA AAT CGC ATG AAC CTA CGC AGT GAA TGT CCT GTA CCT AAG CGG GTA	40
	T F T S Q N P I N L P S E S P V P K P V	
121	CIT GAT ACA AAT GGT AAA AAA CTG AAT CGT AAT TGG AGT TAT GGC ATT AIT TCC ACA TTT	60
	L D T N G K K L N P N S S Y R I I S T F	
181	TGG GGT GGG TTA GCT GGT GAT GTG TAG CTA GGA AAG TCG CGA AAT TCA GAT GCC CCT TGT	80
	W G A L G G D V Y L G K S P N S D A P C	
241	GCA GAT GGC GTA TTC CGT TAC AAT TCC GAT GTT GGA CCT AGC GGT ACA CCC GTT AGA TTC	100
	P D G V F R Y N S D V G P S G T <u>P V *R F</u>	
301	ATT CCT TTA TCT GGA GCA AAT ATC TTT GAA GAT CAA CTA CTT AAC ATA CAA TTC AAT AIT	120
	I P L S G A N I F E D Q L L N I Q F N I	
361	GCA ACA GTG AAA TTG TGT GGT AGT TAT ACA ATT TGG AAA GTG GGT AAT ATA AAT CGA CAT	140
	P T V K L C V S Y T I W K V G N I N A H	
421	CTA AGG ACG ATG TTG TTG GAG ACG GGA GGA ACC ATA GGG CAA GGA GAT AGC AGC TAT TTC	160
	L R T M L L E T G C T I G Q A D S S Y F	
481	AAG ATT GTT AAA TCA TCA AAA TTT GGT TAC AAC TTA TTG TAT TGC GGT CTT AGT CGC GAT	180
	K I V K S S K F G Y N L L Y C F L T R H	
541	TTT CIT TGT CCA TTT TGT GGT GAT GAT AAC TTC TGT GCA AAA GTG GGT GTA GTT ATT GAA	200
	F L C P F C R D N F C A K V G V V I Q	
601	AAT GGA AAA ACG CCT TTG GCT CTT GTC AAC GAA AAT CCT CTT GAT GTC TTA TTC CAG GAA	220
	N G K R R L A L V N E N P L D V L F Q E	
661	GTC TAG TAACAAATAATGCCCTGCACTAGACTATAGTATGTTTAAAGCCCTGCTGGCTAGCTACTATGTTATTT	221
	V * <u>AAAAAAAAAAAAAAAAAAAA</u>	
738	GTAAATAAACCCTGCTAAGGTATAACTATATATATTTAGCATGGCTTCTTAATAAATTGCTTTCCTTATCGTTT	
817	AAAAAAAAAAAAAAAAAAAA	

**Figure 1.** Primary nucleotide sequence of the 22-kD cathepsin D inhibitor of potato. The deduced amino acid sequence is indicated under the nucleotide sequence. The triangle indicates the cleavage site for the signal peptide, and the first asterisk, at Arg<sup>99</sup>, marks the trypsin-binding reactive site. The second asterisk, at nucleotides 664 to 666, marks the stop codon, and putative polyadenylation signals are underlined.

Received July 13, 1992; accepted September 23, 1992.  
Copyright Clearance Center: 0032-0889/93/101/0703/02.  
The GenBank accession number for the nucleotide sequence reported  
in this article is M96257.

#### LITERATURE CITED

- Bryant J, Green TR, Gurusaddaiah T, Ryan CA** (1976) Proteinase inhibitor II from potatoes: isolation and characterization of its protomer components. *Biochemistry* **15**: 3418–3424
- Hannapel DJ** (1991) Distribution of potato tuber proteins during development. *Am Potato J* **68**: 179–190
- Mares M, Meloun B, Pavlik M, Kostka V, Baudys M** (1989) Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. *FEBS Lett* **251**: 94–98
- Melville JC, Ryan CA** (1972) Chymotrypsin inhibitor I from potatoes: large scale preparation and characterization of its subunit components. *J Biol Chem* **247**: 3445–3453
- Ritonja A, Krizaj I, Mesko P, Kopitar M, Lucovnik P, Strukelj B, Pungercar J, Buttle DJ, Barrett AJ, Turk V** (1990) The amino acid sequence of a novel inhibitor of cathepsin D from potato. *FEBS Lett* **267**: 13–15
- Strukelj B, Pungercar J, Ritonja A, Krizaj I, Gubensek F, Kergar I, Turk V** (1990) Nucleotide and deduced amino acid sequence of an aspartic proteinase inhibitor homologue from potato tubers (*Solanum tuberosum* L.). *Nucleic Acids Res* **18**: 4605
- Suh SG, Peterson JE, Stiekema WJ, Hannapel DJ** (1990) Purification and characterization of the 22-kilodalton potato tuber proteins. *Plant Physiol* **94**: 40–45
- Suh SG, Stiekema WJ, Hannapel DJ** (1991) Proteinase-inhibitor activity and wound-inducible gene expression of the 22-KDa potato-tuber proteins. *Planta* **184**: 423–430