식물학회지 제32권 제2호
Korean J. Botany 32(2): 67-78(1989)

Nucleotide Sequence of a Proteinase Inhibitor I Gene in Potato

Lee, Jong Seob and Jung-Sook Park

감자에 존재하는 단백질분해효소 억제제 I 유전자의 염기서열

李 鍾 燮·朴 廷 淑
(서울大學校 自然科學大學 植物學科)

ABSTRACT

Hybridization of DNA isolated from leaves of Russet Burbank potato with tomato cDNA as a probe revealed the presence of about ten inhibitor I genes in the genome. Screening of a genomic library of Russet Burbank potato resulted in isolation of seven different genomic clones carrying inhibitor I genes. One of the genomic clones, clone 2, contained two EcoRI fragments of 3.4 and 1.8 kb in size, respectively, which were hybridized with the probe. The nucleotide sequence of parts of the hybridizing EcoRI fragments revealed that they contain a complete gene which codes for an open reading frame of 107 amino acids. It is interrupted by two intervening sequences of 502 and 493 bp, situated at the positions of codons 17 and 43, respectively, of the open reading frame. Putative regulatory sequences, TATAAA and CCACT, were found at the 5′ flanking region. In addition, a copy of a 100 bp repeat found at a tomato inhibitor I gene was identified.

INTRODUCTION

In potato, several proteinase inhibitors have been identified and studied in soluble proteins of tubers. As many as thirteen different species of inhibitors are thought to be in potato tubers, representing 15 to 25% of the soluble proteins (Belitz *et al.*, 1971). Among them, at least five different inhibitors are heat-stable, three of which have been purified and characterized extensively (Ryan *et al.*, 1976). They are inhibitor I, inhibitor II and carboxypeptidase inhibitor (CPI).

Inhibitor I proteins with a specificity toward chymotrypsin are a mixture of heterogenous pentamers of 40,000 daltons composed of subunits of 8,000 daltons which consist of two major and two minor protomers (Melville and Ryan, 1972). Inhibitor I represents about 2.5% of soluble proteins in tubers of Russet Burbank potato (Ryan *et al.*, 1976).

As tubers develop, inhibitor I proteins are synthesized and accumulate along with inhibitor II

proteins. Furthermore, inhibitor I was found to be induced to accumulate in leaves when damaged by chewing insects or mechanical wounding. Wound induction of inhibitor I genes was also observed in leaves of tomato (Green and Ryan, 1972; Plunkett *et al.*, 1982). The induction of the inhibitor I genes in leaves of tomato and potato is considered to be mediated systemically by a putative wound signal called the proteinase inhibitor-inducing factor (PIIF), which turned out to be oligosaccharides fragmented from leaf cell walls during injury (Bishop *et al.*, 1981). Inhibitor I is synthesized as a prepro-protein, which is post-translationally processed and compartmentalized into the central vacuole (Shumway *et al.*, 1976; Nelson and Ryan, 1980).

In tobacco, however, inhibitor I accumulates in leaves when placed in an environment of complete darkness for several days (Kuo *et al.*, 1984). Inhibitor I was known to be homologous with two isoinhibitors from barley seeds (Svendsen *et al.*, 1980) and an inhibitor from the leech (Seemuller *et al.*, 1981).

In order to understand the mechanisms by which expression of inhibitor I genes is regulated differentially in various solanaceous plants, information on the structure of inhibitor I genes is essential. It will provide a chance for the identification of regulatory sequences involved in their differential regulations. Also, the promoter identified will be used for further study by plant transformation. Therefore, we determined the copy number of inhibitor I genes in the genome of potato and isolated the genes from a genomic library. One of them was characterized at the nucleotide level in this report.

MATERIALS AND METHODS

Materials. Potato (*Solanum tuberosum* cv. Russet Burbank) was used as plant material and was grown in a green house. *E. coli* strain K802 was used as the host of bacteriophage and JM101 was used for cloning of DNA fragments.

Restriction enzymes, Erase-a-base system, DNA sequencing kit, and nick-translation system were purchased from Promega and used as indicated by the manufacturer. Radioisotopes and *GeneScreen Plus membrane* were purchased from New England Nuclear and nitrocellulose filter was from Fischer Scientific. Other chemicals were from Sigma Chemical Co.

DNA isolations. Plasmid DNA was isolated from *E. coli* as described by Brush *et al.* (1985). Genomic DNA from leaves of potato was isolated by the method of Dellaporta *et al.* (1984). Phage DNA was isolated by the methods of Blattner *et al.* (1977) and Maniatis *et al.* (1982).

Screening of a genomic library. About 5×10^5 bacteriophage were screened by the method of Woo (1979) from a EcoI-partial genomic library constructed with DNAs of Russet Burbank potato which was a gift of D.M. Anderson of Phytogen Corporation, Pasadena, CA. *E. coli* strain K802 was used as the host and nick-translated inserts of tomato inhibitor I cDNA clone, pT_I-24 (Graham *et al.*, 1985) was used as the probe.

Southern hybridization. DNAs were digested with various restriction enzymes, elec-

trophoresed in agarose gels and transferred onto GeneScreen *Plus* by the method of Southern (1975). Pretreatment, hybridization and washing of filters were carried out as described by Wahl *et al.* (1979).

Molecular cloning. EcoRI fragments of genomic clones were subcloned into pUC19 (Yanisch-Perron *et al.*, 1985) as described by Maniatis *et al.* (1982).

DNA sequencing. The nucleotide sequence of an inhibitor I gene was determined by the didcoxynucleotide chain termination method (Sanger *et al.*, 1977) in conjunction with universal primers. Denatured plasmid DNAs were used as templates for DNA synthesis after unidirectional deletions with exonuclease Ⅲ and S1 nuclease by the protocol of Henikoff (1984).

RESULTS

Presence of inhibitor I genes as a multigene family in the potato genome. In order to determine the copy number of inhibitor I genes in the potato genome, genomic DNAs digested with various restriction enzymes were hybridized with the insert of a tomato cDNA clone, pT$_{I}$-24 (Graham *et al.*, 1985). As shown in Fig. 1, multiple fragments were found to contain inhibitor I

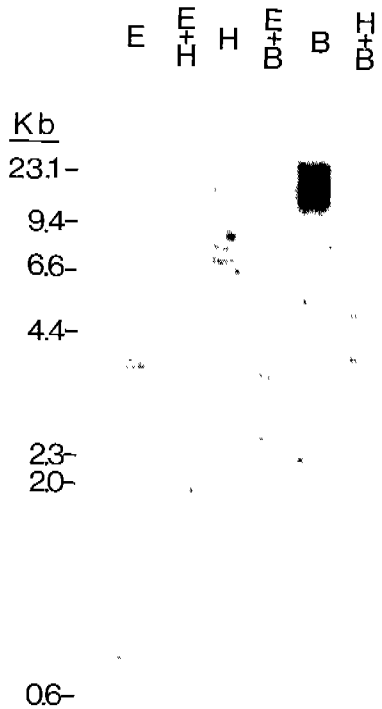

Fig. 1. Determination of the copy number of inhibitor I genes on the genome of Russet Burbank potato. DNA from leaves was hybridized with radioactive inserts of a tomato inhibitor I cDNA clone (Graham *et al.*, 1985) after digestions with various restriction enzymes. Symbols: E, EcoRI; H, HindⅢ; B, BamHI.

genes in the genome of Russet Burbank potato. EcoRI fragments hybridizing with the probe are 9.3, 8.4, 7.5, 6.9, 5.8, 3.7, 3.0, 2.5, 2.3, 1.8, 1.5 and 0.7 kb in size, respectively. The 3.7 kb EcoRI fragment may represent multiple bands as judged by its intensity.

Isolation of inhibitor I genes from a genomic library. Screening of an EcoRI-partial genomic library of Russet Burbank potato gave rise to seven different clones designated as clones 2, 6, 11, 15, 25 and 28 which were hybridized with the tomato inhibitor I cDNA. DNAs isolated from these clones were subjected to Southern hybridization after digestion with EcoRI. It revealed the presence of two EcoRI fragments of 3.4 and 1.8 kb in size in clones 2 and 25, one framgent of 2.5 kb on clone 6, one fragment of 2.4 kb on clones 11 and 16, two fragments of 6.9 and 5.8 kb on clone 15, and one frament of 0.7 kb on clone 28, respectively (data not shown). Of these clones, clone 2 was chosen for subcloning and sequencing of the EcoRI fragments.

Primary structure of an inhibitor I gene in clone 2. The two EcoRI fragments on clone 2 which were hybridized with the probe were subcloned into the EcoRIsite of pUC19. Southern hybridizations of plasmid DNAs of the subclones with a 5′- or 3′-specific probe revealed that the 3.4 kb EcoRI fragment contains the 5′ region of an inhibitor I gene while the 1.8 kb fragment contains the 3′ region (data not shown). DNA sequencing of both ends of each framgent further localized the exact positions of the inhibitor I sequences on the EcoRI fragments. On the basis of these results, the EcoRI fragments were subjected to unidirectional deletions with exonuclease Ⅲ and S1 nuclease, folllowed by DNA sequencing as shown in Fig. 2. The nucleotide sequence of part of the two EcoRI fragments on clone 2 is shown in Fig. 3.

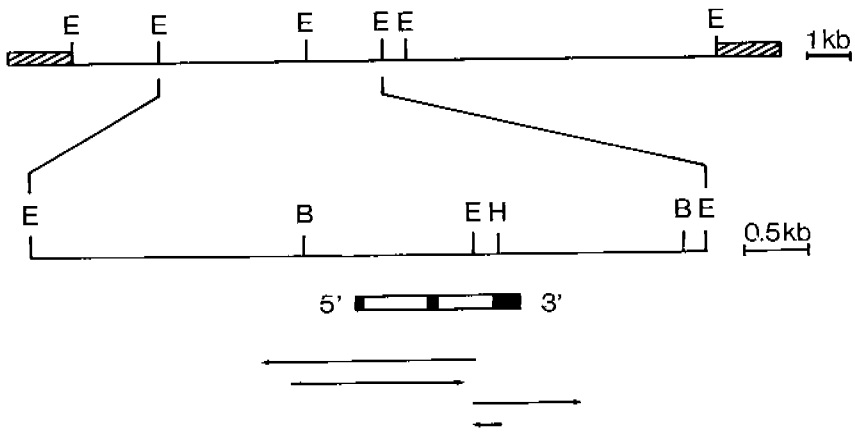

Fig. 2. Restriction map of clone 2 carrying a potato inhibitor I gene and the strategy for determining the nucleotide sequence. Two EcoRI fragments of 3.4 and 1.8 kb in size hybridizing with the probe were subcloned separately into the EcoRI site of pUC19 and subjected to exonulcase Ⅲ and S1 nuclease deletions as described by Henikoff (1984) prior to sequencing. Horizontal arrows indicate the extent and direction of sequencings. Solid boxes indicate exons and open boxes indicate introns. Symbols: B, BglⅡ; E, EcoRI; H, HindⅢ.

```
                                                                   50                                                  100
ATCTTGGCTATATATGAATTTAGTATCATGTAGTGCTTATGTTATTTAGGTACTTTGGCTTCCATTGAAAGCTAGCTTATGTTAGTTAGGTACTTTTTCA
                                                                  150                                                  200
AGTAATTTTGAATGTAATTTTTATTGCTTATTTTTATTATTTTGGTAGATTGTCAAATGTCAACCACTAAAGCTCGACCATGGCAGTCGGTAAGAACAAGC
                                                                  250                                                  300
GAATTTCCAAAGGAAAGAAGGGTCTTGGGCTTTTAAAAAACCACTTACAATTTTAATTTTCTTTAAAAAATGAGCTTACAAATCCAACATGTATTTTATA
                                                                  350                                                  400
TATTTATAGATCTTCTATGTTATTTTGTAACAATCTCCTTTAAAATAAAATTAAAAAAACCTATATACAACTAGAAAATTTATAAACGACTATAAAATTA
                                                                  450                                                  500
AATCTCCCGCATTTACCCTTATCCAATTTATAATATACCGATTAGTGGTATTAGTGTATATGATGTGTCCCCCACACACAATACGTGAATTAAATTTTCA
                                                                  550                                                  600
CTGTCCCTCCTCACTTCTCGTATGTAATTAAAAATATATCAAAACCTAATAATTTTTTCCTTTATCTTCAAATTGTTCAACATGTGGTTGTTAGACACCA
                                                                  650                                                  700
CTAGTAAATAGTACATCACTAGTCACTACAATGAAGGCAACCTGTGCGTATAAATTTATGTGATGCACTCATACAAATTCACTCAATTCCTTCTACTCTT
                                                                  750
TACAACTAAAAGAAA ATG GAG TTA AAG TTT GCT CAC ATC ATT GTT TTC TTT CTT CTT GCA ACT T/ GTTAGTACCCCCCTCCT
                Met Glu Leu Lys Phe Ala His Ile Ile Val Phe Phe Leu Leu Ala Thr
                800                                                  850
CTCTCTCTCTCTCTCTCTCTCTCTATATATATATATATATATATATATATATATATATATTGCAAGAATATTATTAGAGTTGTAAATGATTTATATTTCT
                900                                                  950
TTTACGATATGGGGGTTAGGGTGTGTGTGGGGGGGAGGTCACAAAGGATTTACCTTATACTATATATACATAAAAAAAAAAAGTTTGATCTTGTAAAT
               1000                                                 1050
ATAATGTGACTAACCGGCAAAACAGGTTCGAAAGGGAGACTTGTTCAATGACTTTGTCGATTGATATCCTTTTGCTGAAAAATTACATTATTTAGTGAAA
               1100                                                 1150
ATGTTGACTCTCTTAGTGAAAATTATATATTGACTCTCTTAGTGATAATGTTGGCTCTCCAGAGTTCGGATGAATCTCCTTCTACCCTTGCTTGTGCCAA
               1200                                                 1250
AAAATTTTAGGCGTCATAATTTTCTGATAATCTAACTTGAAATGAGAAATACACTAATTGCTATTGAAATGTTTTTTTTTGGCAG/ CC TTT GAA
                                                                                                  Ser Phe Glu
                                1300                                                                 1350
ACT CTC ATG GCA CGA AAA GAA AGC GAT GGA CCA GAA GTC ATA CAA CTT CTA AAG GAA TTT CAA TGC AAA G/ GTAAA
Thr Leu Met Ala Arg Lys Glu Ser Asp Gly Pro Glu Val Ile Gln Leu Leu Lys Glu Phe Gln Cys Asn
               1400                                                 1450
GTTCTTTTAAAGGAATTTGCATTTTTAGCTTCTCAAAAAAGAGCTTTTTTCTACTATTAAAAACACTTCTTTTTTTTCTTAAATCCTTGGCCAAACACTTC
               1500                                                 1550
AATTCTCTAAAATATACTTTTTTAAATAAGAAGTCATTGTTGACTTTTCTGAGATTTGGCCAAACAAGATGTAGACTATTCAACTCTGGAAAATTTTATT
               1600                                                 1650
AGTTCAGTTGGTTAATTACCTAAAATTTTACATTATTAATAACGATCGAATTCTCCACTTAAAATCTCTTCCGCTACTCCCACTTTTTAAAAAAAAGTAT
               1700                                                 1750
ATATCATGATTCAACTCAATATCTGCTATTATAATAGCACATATAGAGTATATTATTCTTGTTCGTAATTTCTTTCTCTCATTATAATTAATTATTATGT
                                                1800
GTATATGTAG/ GA AAA CTA AGG TGG CCA GAA CTT ATT GGT GTA CCA ACA AAG CTT GCT AAG GGG ATA ATT GAG AAG
          Gly Lys Glu Arg Trp Pro Glu Leu Ile Gly Val Pro Thr Lys Leu Ala Lys Gly Ile Ile Glu Lys
                1850                                                                 1900
GAA AAT TCA CTC ATA AGT AAT GTT CAT ATA TTA TTG AAT GGT TCT CCA GTC ACA TTG GAT ATT CGT TGT GAT CGA
Glu Asn Ser Leu Ile Ser Asn Val His Ile Leu Leu Asn Gly Ser Pro Val Thr Leu Asp Ile Arg Cys Asp Arg
                                1950
GTT CGT CTT TTT GAT AAC ATC TTG GGT TAT GTT GTA GAC ATA CCT GTG GTT GGT TAA TTAATGGATTAATATGGAAGTAAT
Val Arg Leu Phe Asp Asn Ile Leu Gly Tyr Val Val Asp Ile Pro Val Val Gly ***
               2000                                                 2050
TAAGCAGCCACATGTCAAAAATAATAAGGGTTAATGTCTGTTGATTATAATGTCTCCATGTACTCTTACTATAAATAATGAAATAAATAAGTGTGGCTTA
               2100                                                 2150
ATTAAATCCTTTTACATATTGTGCATTACTTTGGCTGTTTAAGTTTGTTTGTATGATTTTGATTCGACATAATTCCAAAAAGTAAACTAGGGTTTTAATT
               2200                                                 2250
CTTGTGGTTGAATAGAAAATATGTGGAATGTATCAAAAATGATATTTAATCTTGTAATTTTAAACATGTCAGGTAGTAAATTAAAATTAAAAGAGTTATC
               2300                                                 2350
ATAAAAGAAAAAAAAAATTCTAAACAAAGTAAGAAATAAAGTAAACAAACAAATTAAAACGCAGAGAGTATTTGACTTAATTGAATAATCTCATGACCAA
               2400
ACGGCTAAATGTCATGGGAGTCTTTCTTCGAAGTATTCAAATTTATTATTTATCAAT
```

Fig. 3. Nucleotide sequence of the inhibitor I gene. DNA sequences at the 5′ flanking region which resemble signals for transcriptional regulation of other eukaryotic genes are boxed. *** is stop codon. A signal for polyadenylation is also boxed at the 3′ untranslational region. A sequence of about 100 bp in size homologous to the direct repeat found at the 5′ flanking region of a tomato inhibitor I gene is underlined (see Fig. 4 for more details).

It indicated that the two EcoRI fragments contain an inhibitor I gene with an open reading frame of 107 amino acids interrupted by two intervening sequences of 502 and 492 bp in length, respectively, as compared with that of tomato inhibitor I gene (Graham *et al.*, 1985; Lee *et al.*, 1986). They are at the positions of codons 17 and 43 and are flanked with GT and AG, as typical of other eukaryotic genes (Brown, 1984).

The region 5′ from the initiation codon (ATG) includes two possible regulatory sequences, TATAAA and CCACT, common to other eukaryotic genes (Breathnach and Chambon, 1981). The 3′ non-coding region contains the sequence, AATAAA, a typical sequence for poly(A) addition (Wickens and Stephensen, 1984). The transcription start site was assigned to be "A" at residue 695 as compared with that of a tomato inhibitor I gene (Lee *et al.*, 1986). Thus, these results suggest that this inhibitor I gene may be a complete and functional gene.

The 5′ flanking region of an inhibitor I gene of tomato revealed the presence of a direct repeat of about 100 bp long which may be involved in wound induction (Lee *et al.*, 1986). Therefore, the 5′ flanking region of potato inhibitor I gene was searched for the presence of the repeat. It revealed a copy of the repeat at the same region as R1 in tomato gene with about 90% homology. Another inhibitor I gene on a 3.7 EcoRI genomic fragment was characterized at the nucleotide level previously (Cleveland *et al.*, 1987). The potato inhibitor I gene was also found to contain a copy of the repeat. Fig. 4 shows the comparison of the nucleotide sequence of the repeat found in the inhibitor I genes of tomato and potato. The repeats include homology to the core nucleotide sequence of enhancer elements, GTGGTTG (Laimins *et al.*, 1983) in addition to the TATA and CAT boxes.

Amino acid sequence of inhibitor I prepro-proteins. As shown in Fig. 5, the amino acid

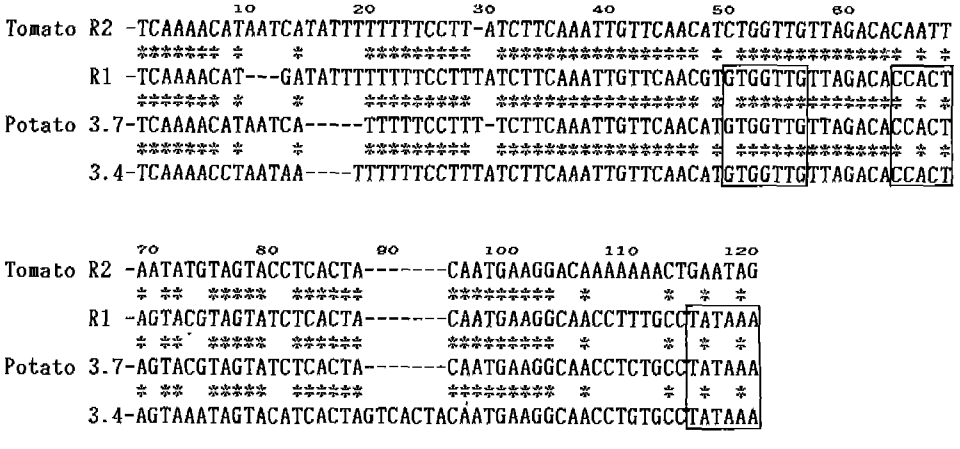

Fig. 4. Comparison of the ~100 bp repeat identified at the inhibitor I genes of tomato (RI proximal to the transcription start site and R2 present 440 bp upstream of the RI) and potato (3.7 on the 3.7 kb and 3.4 on the 3.4 kb fragments). Astrisks denote homology among the repeats. Enhancer core, CCCACT and TATAAA sequences are boxed.

```
                                                    10                                  20
Potato 1 :  Met Glu Leu Lys Phe Ala His Ile Ile Val Phe Phe Leu Leu Ala Thr Ser Phe Glu Thr
Potato 2 :  Met Glu Ser Lys Phe Ala His Ile Ile Val Phe Phe Leu Leu Ala Thr Ser Phe Glu Thr
Tomato 1 :  Met Glu Ser Lys Phe Ala His Ile Ile Val Phe Phe Leu Leu Ala Thr Ser Phe Glu Thr

                                                    30                                  40
Potato 1 :  Leu Met Ala↓Arg Lys Glu Ser Asp Gly Pro Glu Val Ile Gln Leu Leu↑Lys Glu Phe Gln
Potato 2 :  Leu Leu Ala↓Arg Lys Glu Ser Asp Gly Pro Glu Val Ile Glu Leu Gln↑Lys Glu Phe Glu
Tomato 1 :  Leu Met Ala↓Arg Lys Glu Ile Asp Gly Pro Glu Val Ile Glu Leu Leu Lys Glu Phe Asp

                                                    50                                  60
Potato 1 :  --- --- --- --- Cys Asn Gly Lys Glu Arg Trp Pro Glu Leu Ile Gly Val Pro Thr Lys
Potato 2 :  --- --- --- --- Cys Asn Gly Lys Glu Arg Trp Pro Glu Leu Ile Gly Val Pro Thr Lys
Tomato 1 :  Ser Asn↑Leu Met Cys Glu Gly Lys Gln Met Trp Pro Glu Leu Ile Gly Val Pro Thr Lys

                                                    70                                  80
Potato 1 :  Leu Ala Lys Gly Ile Ile Glu Lys Glu Asn Ser Leu Ile Ser Asn Val His Ile Leu Leu
Potato 2 :  Leu Ala Lys Gly Ile Ile Glu Lys Glu Asn Ser Leu Ile Thr Asn Val Gln Ile Leu Leu
Tomato 1 :  Leu Ala Lys Glu Ile Ile Glu Lys Glu Asn Pro Ser Ile Thr Asn Ile Pro Ile Leu Leu

                                                    90                                  100
Potato 1 :  Asn Gly Ser Pro Val Thr [Leu Asp] Ile Arg Cys Asp Arg Val Arg Leu Phe Asp Asn Ile
Potato 2 :  Asn Gly Ser Pro Val Thr [Met Asp] Tyr Arg Cys Asn Arg Val Arg Leu Phe Asp Asn Ile
Tomato 1 :  Ser Gly Ser Pro Ile Thr [Leu Asp] Tyr Leu Cys Asp Arg Val Arg Leu Phe Asp Asn Ile

                                                    110
Potato 1 :  Leu Gly Tyr Val Val Asp Ile Pro Val Val Gly
Potato 2 :  Leu Gly Asp Val Val Gln Ile Pro Arg Val Ala
Tomato 1 :  Leu Gly Phe Val Val Gln Met Pro Val Val Thr
```

Fig. 5. Comparison of the amino acid sequences of inhibitor I prepro-proteins deduced from the DNA sequences of their genes. Regions of difference are underlined. Potato 1 is the amino acid sequence deduced from the nucleotide sequnce of the inhibitor I gene present on 3.4 and 1.8 kb EcoRI fragments. Potato 2 is that of from the inhibitor I gene on a 3.7 kb EcoRI fragment (Cleveland *et al.*, 1987). Tomato 1 is that from a tomato inhibitor I gene (Lee *et al.*, 1986). The reactive sites for chymotrypsin are boxed. The cleavage sites of preproteins are indicated by arrows headed below (↓) and those of proproteins are indicated by arrows headed above (↑).

sequence deduced from the nucleotide sequence of the inhibitor I gene was compared with those of inhibitor I genes of tomato and potato (Lee *et al.*, 1986; Cleveland *et al.*, 1987). Inhibitor I proteins of potato as well as tomato are synthesized as a prepro-protein. Potato pre-inhibitor I is short of four amino acids from that of tomato (Ser-Asn-Leu-Met). Since the mature inhibitor I of potato exhibits the N-terminal Lys-Glu-Phe (Richardson and Cossins, 1975), it indicates that 36 out of 107 amino acids are lost as a result of post-translational processing, representing 34% of the original molecule. The signal sequence or transit peptide is composed of 23 amino acids and exhibits hydrophobicity. It will be apparently cleaved before or during transport into the central vacuole (Walker-Simmons and Ryan, 1977). The prepeptide is composed of 13

amino acids which may be processed by a proteinase in the central vecuole. The reactive (inhibitory) site on the inhibitor I synthesized from this gene was identified as Leu-Asp whch is the same with that of tomato while the reactive site of the other potato gene is Met-Asp (boxed in Fig. 5). Two potato inhibitor I prepro-proteins share homology of 87% from each other while each one shares 80% with tomato inhibitor I.

DISCUSSION

Proteinase inhibitors are usually found in seeds and tubers of plants (Laskowski and Kato, 1980). Proteinases that are inhibited by plant inhibitor proteins are serine proteinases such as chymotrypsin and trypsin. Since proteinases inhibited by plant inhibitor proteins are found to be present in fluids or secretions of animals and microorganisms, plants are thought to synthesize inhibitor proteins as defensive chemicals (Ryan, 1981).

Inhibitor I proteins are present as multiple forms in tubers, suggesting that they may be encoded by a family of related sequences. Genomic hybridization shown in Fig. 1 clearly indicates that potato inhibitor I genes compose one of multigene famlies in plants (Lee, 1988). Since four protomers of inhibitor I proteins are different from each other in amino acid composition, they are considered to be encoded by different genes. About ten EcoRI fragments were hybridized with tomato inhibitor I cDNA (Fig. 1). Each fragment may contain a complete or part of an inhibitor I gene. It is possible that some of inhibitor I sequences are malfunctional as pseudogenes. One possible explanation for the presence of several different inhibitor genes in the potato genome is that different members of this gene family are regulated by different environmental and developmental signals.

Screening of a genomic library of Russet Burbank potato resulted in isolation of EcoRI fragments containing inhibitor I sequences. When compared with those identified by genomic hybridization, the EcoRI fragments of genomic clones were included in the genomic fragments, indicating that they are indeed genomic fragments containing inhibitor I genes. At least five different genes were isolated from the genomic library.

The nucleotide sequence of an inhibitor I gene showed that it contains all the putative regulatory sequences, TATA and CAT boxes and a polyadenylation signal (Fig. 3). Furthermore, the 5′ flanking region upto about 500 nucleotides from the initiation codon (ATG) of this potato inhibitor I gene showed homology of over 80% with inhibitor I genes of tomato and potato previously characterized, It contains a copy of a direct repeat that is found at the same region of tomato inhibitor I gene (Fig. 4). These suggest that the inhibitor I gene reported here may be wound-inducible.

Tandomly repeated promoter elements have been commonly found in various eukaryotic genes. Multiple copies of homologous elements have been shown to be required for full transcriptional activity, and in some cases bind to specific transcription factors (McKnight and Tjian, 1986; Maniatis *et al.*, 1987). The difference in the number of the repeat between tomato

and potato inhibitor I genes may reflect the mode of their differential expression. It is known that tomato inhibitor I genes are under strict wound induction while potato genes are expressed at low level in leaves without wounding. Patatin genes also showed the presence of a long direct repeat at the 5′ flanking region (Rocha-Sosa *et al.*, 1989). The 100 bp repeat found at the 5′ flanking region of inhibitor I genes also contains a short sequence homologous to eukaryotic and viral enhancer elements (Laimins *et al.*, 1983). A similar sequence present at the 5′ flanking region of the *rbcS-3A* gene coding for the small subunit of ribulose-bisphosphate carboxylase in pea is known to direct light-regulated and cell-specific expression in transgenic tobacco (Aoyagi *et al.*, 1988; Kuhlemeier *et al.*, 1988). The functionality of the repeat found at the 5′ flanking region of the inhibitor I genes remains to be investigated by gene transformation experiments with deletions of the region.

The amino acid sequences deduced from the coding regions of two inhibitor I genes so far identified in potato were compared with that of a tomato gene (Fig. 5). It revealed that heterogeneity between potato genes was found to be at the positions at which inhibitor I protomers are varied (Richardson and Cossins. 1975). It indicates that they code for different protomers. The two potato genes showed the same level of divergence at the amino acid level with the tomato gene, reflecting that they might have been in concerted evolution. It is presumed that conversion of proproteins to the mature inhibitor I is mediated by a proteinase present in the central vacuole of potato and tomato plants. The cleavage site in potato proproteins is Leu-X or Gln-X but that in tomato proprotein is Asn-X (Fig. 5). The insertion of four amino acids in the tomato proprotein may have provided a new cleavage site in the tomato.

Inhibitor I genes are found to be regulated in different manners in various plants. Inhibitor I genes are under developmental control in potato tubers while they are wound-inducible in leaves of tomato and potato plants. In tobacco leaves, they are expressed during senescence. It was found to be present in seeds of broad bean and barley and in the leech (Seemuller *et al.*, 1980). Therefore, inhibitor I is distributed widely in nature. It suggests that inhibitor I genes may have been generated before plant and animal diverged about 1 billion years ago. Characterization of inhibitor I genes of various organisms at the nucleotide level may help to understanding their evolutionary relationships.

ACKNOWLEGEMENTS

We thank Dr. C.A. Ryan for giving the opportunity to initiate this work at the Institute of Biological Chemistry, Washington State University, WA, USA. The present investigation was supported by a grant (#87214) from Korea Science and Engineering Foundation.

적 요

감자의 잎에서 DNA를 추출하여 혼성화반응를 수행한 결과 감자의 genome에는 단백질분해효소 억제제

I 유전자가 약 10개가 존재하고 있음을 확인하였다. 감자의 유전자은행에서 7개의 λ clone을 수확하여 이들 중에서 clone 2를 대상으로 억제제 I 유전자를 간직하고 있는 EcoRI절편을 pUC19에 cloning한 후 염기서열을 결정하였다. 그 결과 크기가 3.4kb와 1.8kb인 두 EcoRI 절편에 하나의 억제제 I 유전자가 존재하고 있었으며 이 유전자는 2개의 intron에 의해 나뉘어져 있었고 그들의 크기는 502 bp와 493bp이었다. 이 유전자는 107개의 아미노산으로 구성된 억제제 I의 prepro-protein을 암호화하고 있었다. 5′ 근접부위에서 진핵생물의 유전자 전사에 조절역할을 수행하고 있는 것으로 알려진 TATAAA와 CCACT의 염기서열이 발견되었다. 또한, 토마도 억제제 I 유전자에서 발견된 약 100bp로 구성된 반복서열도 존재하고 있음을 확인하였다.

REFERENCES

Aoyagi, K., C. Kuhlemeier and N.-H. Chua. 1988. The pea *rbcS-3A* enhancer-like elements direct cell-specific expression in transgenic tobacco. *Mol. Gen. Genet.* **213**: 179–185.

Belitz, H.D., K.P. Kaiser and K. Santarius. 1971. Trypsin and chymotrypsin inhibitors from potatoes: isolation and some properties. *Biochem. Biophys. Res. Commun.* **42**: 420–427.

Bishop, P.D., D.J. Makus. G. Pearce and C.A. Ryan. 1981. Proteinase inhibitor inducing factor activity in tomato leaves resides in oligosaccharides enzymatically released from cell wall. *Proc. Natl. Acad. Sci. USA* **78**: 3536–3540.

Blattner, F.R., B.G. Williams, A.E. Blechl, K. Denniston-Thompson, H.E. Faber, L.-A. Furlong, D.J. Grunwald, D.O. Kiefer, D.D. More, E.L. Sheldon and O. Smithies. 1977. Charon phages: Safer derivatives of bacteriophage lambda for DNA cloning. *Science* **196**: 161–169.

Breathnach, R. and P. Chambon. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Ann. Rev. Biochem.* **50**: 349–383.

Brown, J.W. 1986. A catalogue of splice junction and putative branch point sequences from plant introns. *Nucl. Acids Res.* **14**: 944–9559.

Brush, D., J.B. Dodgson, O.-R. Choi, P.W. Stevens and J.D. Engel. 1985. Replacement varient histones genes contain intervening sequences. *Mol. Cell. Biol.* **5**: 1307–1317.

Cleveland, T.E., R.W. Thornburg and C.A. Ryan. 1987. Molecular characterizatioon of a wound-inducible inhibitor I gene from potato and the processing of its mRNA and protein. *Plant Mol. Biol.* **8**: 199–207.

Dellaporta, S.L., J. Wood and J.B. Hicks. 1984. Maize DNA miniprep. *In,* Molecular Biolgy of Plants: A laboratory course manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 36–37.

Graham, J.S., G. Pearce, J. Merryweather, K. Titani, L.H. Ericsson and C.A. Ryan. 1985. Wound-induced proteinase inhibitors from tomato leaves. I. The cDNA-deduced sequence of pre-inhibitor I and its post-translational processing. *J. Biol. Chem.* **260**: 6555–6560.

Green, T.R. and C.A. Ryan. 1972. Wound-induced proteinase inhibitors in plant leaves: a possible defense mechanism against insects. *Science* **175**: 776–777.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targetted breakpoints for DNA sequencing. *Gene* **28**: 351–359.

Kuhlemeier, C., M. Cuozzo, P.J. Green, E. Goyvaerts, K. Ward and N.-H. Chua. 1988. Localization and

conditional redundancy of regulatory elements on *rbcS-3A*, a pea gene encoding the small subunit of ribulose-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. USA* **85**: 4662–4666.

Kuo, T.-M., G. Pearce and C.A. Ryan. 1984. Isolation and characterization of proteinase inhibitor I from etiolated tobacco leaves. *Arch. Biochem. Biophys.* **230**: 504–510.

Laimins, L.A., M. Kessel, N. Rosenthal and G. Khoury. 1983. Viral and cellular enhancer elements. *In*, Enhancers and Eukaryotic Gene Expression, Y. Gluzman and T. Shenk (eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. pp. 28–37.

Laskowski, M. and I. Kato. 1980. Protein inhibitors of proteinases. *Ann. Rev. Biochem.* **49**: 593–626.

Lee, J.S. 1988. Molecular characterization of proteinase inhibitor genes in plants. *In*, Proceeding of the Joint Korean-German Symposium on Molecular Genetics, The Genetics Society of Korea, Seoul. pp 71–82.

Lee, J.S., W.E. Brown, J.S. Graham, G. Pearce, E.H. Fox, T.W. Dreher, K.D. Ahern, G.D. Pearson and C.A. Ryan. 1986. Molecular characterization and phylogenetic studies of a wound-inducible proteinase inhibitor I gene in *Lycopersicon* species. *Proc. Natl. Acad. Sci. USA* **83**: 7277–7281.

Maniatis, T., E.F. Fritsch and J. Sambrook. 1982. Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maniatis, T., S. Goodbourn and J.A. Fischer. 1987. Regulation of inducible and tissue-specific gene expression. *Science* **236**: 1237–1245.

McKnight, S. and R. Tjian. 1986. Transcriptional selectivity of viral genes in mammalian cells. *Cell* **46**: 795–805.

Melvile, J.C. and C.A. Ryan. 1972. Chymotrypsin inhibitor I from potatoes: Large scale preparation and the characterization of its subunit components. *J. Biol. Chem.* **247**: 3445–3453.

Nelson, C.E. and C.A. Ryan. 1980. *In vitro* synthesis of pre-proteins or vacuolar compartmented proteinase inhibitors that accumulate in leaves of wounded tomato plants. *Proc. Natl. Acad. Sci. USA* **77**: 1975–1979.

Plunkett, G., D.F. Senear, G. Zuroske and C.A. Ryan. 1982. Proteinase inhibitors I and II from leaves of wounded tomato plants: purification and properties. *Arch. Biochem. Biophyys.* **213**: 463–472.

Richardson, M. and L. Cossins. 1975. Chymotryptic inhibitor I from potatoes: The amino acid sequences of subunits B, C and D. *FEBS Letters* **45**: 11–13.

Rocha-Sosa, M., U. Sonnewald, W. Frommer, M. Stratmann, J. Schell and L. Willmitzer. 1989. Both developmental and metabolic signals activate the promoter of a class I patain gene. *EMBO J.* **8**: 23–29.

Ryan, C.A. 1981. Proteinase inhibitors. *In*, The Biochemistry of Plants, P.K. Stumpf and E.E. Conn (eds.). Vol. 6, Academic Press, New York. pp. 351–370.

Ryan, C.A., T. Kuo, G. Pearce and R. Kunkel. 1976. Variability in the concentration of three heat stable proteinase inhibitor proteins in potato tubers. *Am. Potato J.* **53**: 443–455.

Sanger, F., S. Nicklen and A.R. Couson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463–5467.

Seemuller, U.M, Euiltz, H. Fritz and A. Strobl. 1981. Structure of the elastase-cathpsin G, inhibitor from the leech *Hirudo medicinalis*. *Hoppe-Seyler's Z. Physiol. Chem.* **361**: 1841–1846.

Shumway, L.K., V.V. Yang and C.A. Ryan. 1976. Evidence for the presence of proteinase inhibitor I in vacuolar protein bodies of plant cells. *Planta* **129**: 161–165.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503–517.

Svendsen, I., I. Jonassen, J. Hejgaard and S. Borsen. 1980. Amino acid sequence homology between a serine protease inhibitor from barley and potato inhibitor I. *Carlsberg Res. Comm.* **45**: 393–502.

Wahl, G.M., M. Stern and G.R. Stark. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzloxymetal-paper and rapid hybridization by using dextran sulfate. *Proc. Natl. Acad. Sci. USA* **76**: 3683–3687.

Walker-Simmons, M. and C.A. Ryan. 1977. Immunological identification of proteinase inhibitor I and II in isolated tomato leaf vacuole. *Plant Physiol.* **60**: 61–63.

Wickens, M. and P. Stephenson. 1984. Role of the conserved AAUAAA sequence: four point mutants prevent messenger RNA 3′ end formation. *Science* **226**: 1045–1051.

Woo, S.L.C. 1979. A sensitive and rapid method for recombinant phage screening. *Methods Enzymol.* **68**: 389–395.

Yanisch-Perron, C., J. Vieira and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103–119.

(Received February, 17, 1989)