Volume 17 Number 14 1989 Nucleic Acids Research

Sequence analysis of a genomic clone encoding an endochitinase from *Solanum tuberosum*

John J.Gaynor and Karl M.Unkenholz

Department of Biological Sciences, Rutgers University, Newark, NJ 07102-1814, USA
Submitted May 22, 1989 EMBL accession no. X15494

A genomic clone (pRU8713) encoding an endochitinase from potato (*Solanum tuberosum* L. *cv.* Russet Burbank) has been isolated and sequenced (Figure 1). A comparison of this genomic sequence to a previously isolated cDNA clone for a potato endochitinase (1) show this gene to be intron-less. The coding region of this clone is highly homologous to the endochitinase from *Phaseolus* (46.7% ; 2), *Nicotiana* (73.9% ; 3), and *Lycoperiscon* (85.6% ; 4). A more detailed analysis of this clone and its expression will be published elsewhere (5).

```
AAAAAAATAT ATTATATTGT TTTTAAAATA TATTTATATT AAATATTTTA TTTTAAATAA ATTAACTTAT   70
TAGATAAATA AACGATTATA TAAAAACATA TTTTTTTAAA TTAATCATTT ATAAAACACC AACACATTAT  140
TTCACGAAAA ATATTTATAT CACCACACCA ATATCGACGA ATAATAAAAT ATAGATTTAT TTCGAAAGTA  210
TTTTCATCAT ATAAACACAC ACCAGTATTA TTATTGAGTT TCGATTTTAT TTCACTCGAC CCCCCCTCAC  280
CCGCACCTTC ACCCGCCTTA CACCACACCC CCCCCCCCCC CAAAAAAAAT AATATAGATT TCATTTTAAA  350
TATATAAATT AATTTATTAC ATTAAATATT TATTTTTTTT ACACAAAAAA TATTATTTAT TTAAAAATAT  420
ATTTACATTA AAATATATAT TTTTTCACGC CTCACTCACC TCGACCCACC CCCCACCTCA CCCGTAAAAT  490
AAAGATTTGA TTTGTTTTTC GAAGATTATT TACATTAAAA TATTATTAAA AATAAAATTT CACCCCTCAC  560
CTACACCCCC CCACCCATAT CACCCACCCC ATATTTTTTT TTTGATTTGT TGTTAAAATA TATTTTACAT  630
TAAATATTTA TTCACCTTTC ACCCCCCCCC TCACCAACAC CCCCACCCCC ACCCCCCCCA CCCCCCTGTA  700
CATTGGGTCG ATTGATTTAT CGATATCGAC CCACAAATAT ATATTTATCG CATATCACCC ATTATTTGTG  770
ATTCATCGCG ATATAGATTT ACAACAAAAT AATATAGTGG CGCGCGAGGC GCGCCTCACC CAGTATTACA  840
TCGAATACAA TAAGATTATA TATTTGTAAA TATTATTTAC GATTATCGAA TATATTCGGA TTCGATACAT  910
ATAATATTAT ATTAATATTA ATATTTATAG TATTATTAAA TAAAAACTTT ATCGGTTAAA TTCAAAATAT  980
TATTTATTAC AAATATTAAA ATAAAATAGT TCATTTGTAT ACATCATTAT TGTTCGATTT TGGCGCCTTG 1050
AGTAGCGACA CCTTTAAAAA TAAAAATAAT AGTATTGAGC GGTTAGAGGG GCGCGCATGG ATATCGACGA 1120
CCAACGATCA CATTTAATAA AATATAATAT TGAGTATTTC GCCCCCGACC TCGCCTTATT TATTACGACA 1190
CCCGATCGTC ACCTTTAAAA TTTATAATAT ATCATTTTAT TCGAAGTATT AATATATCGA AAATATTCAC 1260
CGGCGCCTTA TTTATTACGA CATCGATCAT CACCTTTATA TATCATTTTA TTGTGAGTTT AATATATCGA 1330
AAATATTGAT AAATATATTT ATACGAGTTT ATTGGGTTTA AATACGACGA ATACGACCTC GACCTCGACA 1400
CCTCACGAGA GTTAACGAAG AGCGTCGATT CACCAACAAC GACCGATTTA ACAACGAACA AAATAACAAG 1470
AGCGAAACAG TATTCGAAAC ACGTCGATTA GTTTATCACG ACCCGCATCA ATACACCTCA TACGAACACC 1540
GATTACACCT TTCGATTTCA TTTTCACCTA TTTTGGCGTT ATTTGTTTAT CGCGTTTCGA TAATATAACA 1610
ATATATTTAT TTCGATCACA TCGACCTAGT ATTACGAGGG TCGTTCGCAT ATCATCATAT TTTAATCGCG 1680
CGTTATCGTT ATCGATCGGT TCATAGATTT TCACATCAAA TATTTATTCA TTACATTCGA TATATTATAT 1750
TGTGATCAGA TTTTTTATTT AGAGGTATAT TTACGAGTAC AAAAAATATA CACGAACATT AGATTACAGA 1820
TCAAATATTG ATCGACGAAG ATTGAGCGTA CGGTCAATCT ACCCGCCGTT ATATTTCATT TACATCGTAT 1890
ATATTGAGCG TCACGCCGAG GGTGGCGCCT TATAAATTGG TAAAGATCAT CGTTCATCAA CACCCATTCG 1960
CCTTAATACG ATTTCATTCC AAAATTAAGA AAGAGGAGCA GGAATATGCG TCGACATAAA GAAGTTAATT 2030
TTGTGGCTTA TTTATTGTTT TCCTTGTTAG TTTTGGTGTC CGCTGCCTTA GCTCAGAACT GTGGTTCCCA 2100
GGGGGGAGGC AAAGCTTGTG CCTCGGGACA ATGTTGCAGC AAATTCGGGT GGTGTGGTAA CACCAATGAC 2170
TATTGTGGTT CCGGCAATTG TCAAAGTCAG TGTCCTGGTG GCGGCCCTGG TCCTGGTCCT GGCGGTGATC 2240
TTGGTTCAGC TATTTCTAAT TCTATGTTTG ATCAAATGTT AAAACATCGT AATGAAAATT CTTGTCAAGG 2310
CAAAAATTTT TACTCCTACA ATGCTTTTAT TAATGCTGCT CGTTCTTTTC CTGGTTTTGG TACTTCTGGT 2380
GATATTAATG CTCGAAAGCG AGAAATTGCG GCCTTTTTTG CTCAGACAAG TCATGAGACC ACTGGTGGAT 2450
GGGCTAGTGC TCCAGATGGT CCTTATGCTT GGGGTTATTG TTTCCTTAGA GAAAGAGGTA ACCCCGGTGA 2520
CTACTGTCCA CCAAGCAGTC AATGGCCTTG TGCACCTGGA AGAAAATACT TCGGACGAGG CCCGATCCAA 2590
ATATCACACA ACTACAACTA CGGGCCATGC GGAAGGGCCA TCGGAGTGGA CCTCTTAAAC AATCCTGATT 2660
TGGTAGCCAC TGACCCAGTC ATCTCATTCA AAACTGCTCT CTGGTTTTGG ATGACCCCCC AATCACCAAA 2730
GCCTTCGTGC CACGACGTCA TCATTGGCAG ATGGAACCCA TCTTCCGCTG ACCGAGCAGC CAATCGACTG 2800
CCTGGATTCG GTGTCATCAC GAACATCATT AATGGGGGCC TAGAATGTGG GCGTGGTACT GACAACAGAG 2870
```



```
TACAAGATCG CATTGGATTT TACAGGAGGT ATTGCAGCAT TTTGGGTGTT ACTCCCGGTG ATAATCTAGA 2940
TTGTGTTAAT CAACGTTGGT TTGGTAATGC TCTATTGGTG GACACTTTGT AATCCCACTC TCGTTCGTTG 3010
CTTAATAAAT TCGGGGATAC TCACTAATAA ATATATCTTT ATGAATTTAT AAAACTGATT TTTTTTTTTT 3080
TTTTTTTGTG TGTGTATTTG GTTGTTTATT GTATTTAATA AAAATATATC GGAAATAAAG TTACTAGTGT 3150
TGCTTTGTTA TGTTCATGTT TCGCAATGTT CTCTCTTTAG CGGTIATATC GAGTGTACGT CGTGTGTAGT 3220
GTTATTATCC ATTGGTACAG GTGAAAAAAT AATGCCTTTA TTCTTTTATT GGGTTGGGCC C          3281
```

Figure 1. The nucleotide sequence of a genomic clone for potato endochitinase (pRU8713). The endochitinase coding region is from 2006 to 2992. The transcription start site is at 1956. A putative CAAT box (1851-1860) and TATA box (1921-1927) have been identified.

Acknowledgements: This work was supported by a grant from NSF (DMB 8704677). This is publication number 76 of the Department of Biological Sciences, Rutgers-Newark.

References:

1. Gaynor, J.J. (1988) *Nuc. Acids Res.*, 16:5210.
2. Broglie, K.E., Gaynor, J.J., and R.M. Broglie. (1986) *Proc. Natl. Acad. Sci., USA* 83:6820-6824.
3. Shinshi, H., Mohnen, D., and F. Meins, Jr. (1987) *Proc. Natl. Acad. Sci., USA* 84:89-93.
4. Durand-Tardif, M. (1986) Ph.D. Thesis, Universite de Paris Sud, Centre d'Orsay. 156 pp.
5. Gaynor, J.J., and K.M. Unkenholz. (1989) Submitted to Plant Molecular Biology.