Nucleotide sequence of an actinidin genomic clone

K.C.Snowden and R.C.Gardner
Centre for Gene Technology, Department of Cellular and Molecular Biology, University of Auckland, Private Bag, Auckland, New Zealand

Submitted October 18, 1990 GenBank accession no. M38422

Actinidin is a cysteine protease which is abundant in the fruit of kiwifruit (*Actinidia deliciosa*) and is encoded by a large gene family (1, 2). A cDNA clone for actinidin (3) was used for high stringency screening of a λGEM™-11 genomic library from the kiwifruit cultivar Hayward.

We present below the nucleotide sequence of a genomic actinidin clone, λKIWI44 (start and stop codons, putative TATA box and polyadenylation signal are underlined). The sequence includes the entire coding region for a 380 amino acid protein that shows >98% identity to three previously described cDNA clones (1, 3), and has >92% identity with the 220 amino acid mature protein (4). The actinidin gene consists of 5 exons separated by introns of 412, 91, 913 and 218 bp. The location of the first intron corresponds with that of the first intron of the cysteine protease pst-cathepsin from *Dictyostelium* (5). No other intron location in actinidin corresponds with those in pst-cathepsin or in barley aleurain (6).

ACKNOWLEDGEMENTS

We thank Craig van Dolleweerd for construction of the library and Jeannette Keeling for assistance with the initial library screening. This research was supported by grants from the NZ Kiwifruit Marketing Board and MAFTech.

REFERENCES

1. Praekelt,U.M., McKee,R.A. and Smith,H. (1988) *Plant Mol. Biol.* **10**, 193–202.
2. Keeling,J., Maxwell,P. and Gardner,R.C. (1990) *Plant Mol. Biol.*, (in press).
3. Podivinsky,E., Forster,R.L.S. and Gardner,R.C. (1989) *Nucl. Acids Res.* **17**, 8363.
4. Baker,E.N. (1980) *J. Mol. Biol.* **141**, 441–484.
5. Datta,S. and Firtel,R.A. (1987) *Mol. Cell. Biol.* **7**, 149–159.
6. Whittier,R.F., Dean,D.A. and Rogers,J.C. (1987) *Nucl. Acids Res.* **15**, 2515–2535.

```
   1 CTGCAGTTCCAACAGAACCTTTAAAAAAAAATTGTGAAAATCATTTTTTCAAATGTCGTAAGACCCCCCCCACCCCCCACGCACCCTATATAAAGGCCACTCTCTCCCTCCACATTCACACA  120
 121 CCTCCAATCCCAATCTTTTTCTTCTAAAATTCAAAAAACGAGAGAGAACAAAAATGGGTTTGCCCAAATCCTTCGTCTCAATGTCTCTCCTCTTCTTCTCCACACTCCTTATACTCTCAT  240
                                                                M  G  L  P  K  S  F  V  S  M  S  L  L  F  F  S  T  L  L  I  L  S  L
 241 TAGCCTTCAACGCCAAAAACTTGACCCAGAGAACAAACGATGAAGTCAAGGCCATGTACGAGTCGTGGCTTATCAAGTATGGCAAATCCTATAACTCGCTGGGTGAGTGGGAGAGGAGAT  360
       A  F  N  A  K  N  L  T  Q  R  T  N  D  E  V  K  A  M  Y  E  S  W  L  I  K  Y  G  K  S  Y  N  S  L  G  E  W  E  R  R  F
 361 TTGAGATTTTCAAGGAAACTTTGAGGTTCATTGACGAGCACAATGCAGACACAAACCGTAGTTACAAGGTGGGCCTGAACCAGTTTGCTGACCTGACTGATGAGGAGTTCCGGTCCACTT  480
       E  I  F  K  E  T  L  R  F  I  D  E  H  N  A  D  T  N  R  S  Y  K  V  G  L  N  Q  F  A  D  L  T  D  E  E  F  R  S  T  Y
 481 ACTTGGGATTTACAAGCGGTTCAAATAAGACGAAGGTGAGCAACCGGTACGAGCCCCGAGTCGGCCAAGTATTGCCGAGTTATGTTGATTGGAGGTCGGCAGGGGCCGTGGTTGACATCA  600
       L  G  F  T  S  G  S  N  K  T  K  V  S  N  R  Y  E  P  R  V  G  Q  V  L  P  S  Y  V  D  W  R  S  A  G  A  V  V  D  I  K
 601 AATCCCAGGGTGAATGTGGTATGTCTGATAATCTCTCTCCTATCATGTTTGTGTCGTCGCTTATGTCCTGTTTGGAGTTGATTTTTTATTTATTTTTTTAATGAACTTCTTTTAATTGTT  720
       S  Q  G  E  C  G
 721 TGATCAAAATTTTATAATTTTTTTTACTTGATGCGAGTAGTTTGAAAGTATAAAATTTTGATTAAAGTTATCTAAATTTGTTGTCTTTAGTTTCTTTTTTTTTAAATTTAAATAAAAAGT  840
 841 TTACATTTTAAAATTTTTCTTGTCAATAGGAATTGATAAGTAAAGCTTTGATTAAAAGGTGGAAAAAAATATAAAATTCAAAAACCTAAGCAACTAAAAAGTTTCATAAAAAAACAAAGT  960
 961 CAAACTAGAGCCTTACTAATGTGTATGTTTTGAAATAATAACTCAATGTCTATTAATCATTGGAATGAAGGGGGTTGTTGGGCTTTTTCGGCCATCGCCACGGTGGAAGGGATCAACAAG 1080
                                                                G  C  W  A  F  S  A  I  A  T  V  E  G  I  N  K
1081 ATAGTGACCGGAGTCTTAATTTCGCTGTCAGAACAAGAACTTATAGATTGCGGTAGGACACAAAACACCAGGGGCTGCAATGGCGGTTACATAACCGACGGGTTTCAGTTCATCATCAAC 1200
       I  V  T  G  V  L  I  S  L  S  E  Q  E  L  I  D  C  G  R  T  Q  N  T  R  G  C  N  G  G  Y  I  T  D  G  F  Q  F  I  I  N
1201 AACGGTGGAATTAATACCGAGGAAAATTATCCCTACACGGCTCAAGATGGTGAATGCAACGTGGAATTAGTAAGTACTCAGTACTTCATCATATGGTACAGGATAATAAAAGCAAGAATA 1320
       N  G  G  I  N  T  E  E  N  Y  P  Y  T  A  Q  D  G  E  C  N  V  E  L
1321 AATTTACATTAATTTTTTTTGGTGATTTTTGTTTATGCAGCAAAATGAAAAGTATGTTACAATTGATACTTATGAAAATGTTCCTTATAACAACGAGTGGGCATTGCAAACAGCAGTGAC 1440
                                                                Q  N  E  K  Y  V  T  I  D  T  Y  E  N  V  P  Y  N  N  E  W  A  L  Q  T  A  V  T
1441 ATACCAACCTGTGAGCGTTGCCCTTGACGCCGCTGGCGACGCGTTCAAACAGTATTCATCGGTAAGATTCCTTTACCTTTACGAATATATAAAGGGCTGACAAGAACAATTGGGTGGATT 1560
       Y  Q  P  V  S  V  A  L  D  A  A  G  D  A  F  K  Q  Y  S  S
1561 GGGTGTTTTTTAGGGTAAAAATATGGAAGACTTTTGGACTTCTTTTTTAGGAATTCAAGCATATGGTTGCTAACTAGTTGGCATGTTTCCTATTATGTAGTTTTTTTATATAACATATTT 1680
1681 TTTTTCCATTTTTGCCTATATTTGTGAAGATGATTGGTAGGTTGAGTTTTTTGTAAGTAGCAATACTAGGAGAGCTCGTGTACTCTCTGGAGCTCTTCCACATAAAAATATATTATTATT 1800
1801 TGCATATGCCCAACGTTTTCCATGCCCGTACGTGTATTCTCCCTAATATCTAAGAGATAATTATCATGTTTAAGCCCTTGGATAGTCTCAAGTGATTATGACAAATGGGAGATATGAGTT 1920
1921 ACAGCCATGATGTTTGAGTAATCCATTTCACGTTTGGTGATCATAGGCCCAAGTGGTGCAAGCTAGTTGAATAAATGGTCTGCATCTTTAGCAAATCACAATATTGCATGTGACTTGGAG 2040
2041 ATGAATACCCACTATATAACCGATTGTGATATATCCTTGTGTTATTCTAAAGAATAGGTTCTTTGCGAGGCTAATATGCTCAAAAATTGAAAATTCCACCATCAACAAGATGACTAGTTT 2160
2161 GATGATACTCACCACACCATATTATTTCAAGATGACTAAATTTTATTATAATTTATTAGGATTTAACTTGTGATACAAGTAACGGTATAAGTACCGATAATACAACCTCATATCAATCCT 2280
2281 ACTTTAAACCTTGCTATGATTCTTTTAAAATTCTTCAAACACAAAGAGGAGTGATGCGGACAAGTTATTCGAAGATAGGGGAAATATATACTTGAGAAATCTAATCCTCTTTGGTTTTCG 2400
2401 AATTCTCTCTTTAGGGCATATTCACCGGACCATGTGGAACAGCAATAGACCATGCTGTTACTATTGTTGGATATGGCACAGAGGGAGGTATCGACTATTGGATAGTGAAAAACTCATGGG 2520
                G  I  F  T  G  P  C  G  T  A  I  D  H  A  V  T  I  V  G  Y  G  T  E  G  G  I  D  Y  W  I  V  K  N  S  W  D
2521 ACACGACGTGGGGAGAGGAAGGCTACATGAGGATCCTGCGTAATGTCGGAGGTGCTGGAACATGTGGAATTGCAACGATGCCATCTTACCCTGTCAAGTACAACAACCAAAATTACCCCG 2640
       T  T  W  G  E  E  G  Y  M  R  I  L  R  N  V  G  G  A  G  T  C  G  I  A  T  M  P  S  Y  P  V  K  Y  N  N  Q  N  Y  P  E
2641 AACCATACTCATCTCTCATAAATCCTCCCGCCTTCTCAATGGTAAAATATATACTTCCTTATCTCTCTTTGAAATTTTTTTTTTTCGTTTTCTTTTTTGTGTTCTCTGATAAAAAAGTAA 2760
       P  Y  S  S  L  I  N  P  P  A  F  S  M
2761 CGAAAAGGATTGAGAACACATCTTCTCCATCAAAAACATCTTTTGAAAAATGAAAACAAGAAAGAGAATTCGAAAACCATACCAATTCCAGAAAGATTAAGATTGAGTTTTAATTTATTT 2880
2881 ACATTCTTTGATGCGACAGAGCAAGGATGGCCCAGTGGGAGTAGAAGATGGACAGAGGTACAGCGCTTAGGAAAGTTTCAATGAGATGAGGAAAGCAGGGAGAGAATGTAGATAAATAAT 3000
                S  K  D  G  P  V  G  V  E  D  G  Q  R  Y  S  A  *
3001 ATTGCATCTGAAAAATAAAAGAACCTTTGCTTTCCAGTTATGGAACTTATGTATTGAATAAAGGTGGCTTCAGGGGATTTGTTTGTCCTCCTGATACTGCCTCTACTTTTTACGTGATAA 3120
3121 TAATATAAGTAGTTGTGCTTTCAATTCCTTCAAATTACCTTTTTACTTATAATGCTAATTTGAACAATTAAAACCCAATAATTTTAATTTGGTTTTAATTTAATCCCTTCGTCAATACCTG 3240
3241 TTAGCAATTTAGACGGAAAGAAAGCATGTGAAAACTCACGTGGCAATTTTTAGGGACGAAAGTGAAACATTGGACAAATTACGAAAACCCAATGGTTCAAGTAGCCTCACTTTCCATTCAA 3360
3361 AAAGGTAAACGGGTATTGACAGATGGACTATATTCAAAC 3399
```