

Nucleotide sequence of the gene for the *Vigna mungo* sulfhydryl-endopeptidase (SH-EP)

Harue Akasofu, Daisuke Yamauchi and Takao Minamikawa
Department of Biology, Tokyo Metropolitan University, Setagaya-ku 158, Japan

Submitted March 6, 1990 EMBL accession no. X51900

A genomic clone for SH-EP was isolated from a *Vigna mungo* genomic library in lambda DASH with the SHE-EP cDNA probe (1). We give here the nucleotide sequence of the gene together with the deduced amino acid sequence. Nucleotides are numbered from the initiation codon. Asterisks are positioned below the first and last residues of the nucleotide sequence of cDNA for SH-EP. Initation and termination codons and the putative polyadenylation signal are underlined.

REFERENCE

1. Akasofu,H. *et al.* (1989) *Nucl. Acids Res.* **17**, 6733.

```
-1676 AAGCTTAACTTTTGAGGCAGAGCTTGTAAATTGTAACAGGTGAGGTAGAAAGACGGAAAGTACTTTTAATAATAAAAGGTTTGAAAAATTAAGAAAAGAAGAAGAAAATA
-1566 TTTTGTGAGTGCACGCGATGGATCTAATCCTTCCATGAAAAAGAATATCAAGAATAACAAAAATTGACAAAATCAGCGAATACTTCACCCAAAAGTCTACACAATAATAA
-1456 ATGCTAAATCACATATAATTTGTGATGCATAACGCATTACGCTATCGTAATCCTTTACAACAAGCAAGAACGTCATCCCAGAATCTCAACTCAAATCAAAACCGTTCATT
-1346 CATAAATAAAAAATATTCTTACATTCTTTTGCAAATAGAACCTTTGCCAAATTGAAATAACAAACTCTAGGTATTTGTCAAATTAACTTACCAACTTCTCGTTATATAAT
-1236 TTTAGATTTATAATCATGTCTATAAATTATTTCTATACACTCTCTCTCAAATTTGACCTTTACATTCTGTGATTTATTTGAACAGAATAAATCACTGTAAAACTAAACAA
-1126 CTCTTTAAAAAAGGTAAATTAGGAAAAGTCGAAATCAATAAATTATAAATCAATCCCTAGAAAACTGCAAGATAATATTCTTACCAAAATCATTTAAATAAATTTGTAAG
-1016 TTTTTTCTTTATACCAATTTTCTGAGACCCAGAGACATTCTTAAATTCATAACAACGGTTTTAAGTATCAGAGTATAACATCTTTGTATAAATAGATTTTTGAACGTTCA
 -906 ATAACTAACACGTCAGTTTTTGTTTCCACGTTGTACGTTTAATAACAATAAATGCGTGAGTTAGATTACTAATCAGAAGTTAGAAGTGTACAAGACTAACTTTATACAGA
 -796 AATATATTGTTTCAGACTGCACTTTATGGTGCGTAGCACCTCAAAACTCTTACCTTTCGCATACATTTTCACACTTCATCCAAACCTTTCGAAAAGTCACTTCCCTTATA
 -686 TTAAAGGACTATGATATATAAAAAAGACTATATGTGTTACTAATTTATTGGTTTGTATATTTGTAATAAATCGTTCCATCAAGAGGAGCTATCACATATTGAGAACAGTAAA
 -576 AAAAAAAAAAAGTTGGTAAAAAAACATTTTCTTATATTATATCATAAAATCAGTTACCATAGTATTTTAGAGTTTTCAGAATAATGCTTCACCCAACTTGCAACTCATTG
 -466 TGCCTCAAAACAGGACGTAACCATGTTACTCACTCTCCTGCACAACCCCTTGTTAAACTGATAGCGTGATCAGCATGCAAGAGAAAGATGATTCTTGAAGCATACGATAA
 -356 CAGATTGAATGTGACAAAAAGTTTGTGTCTCAGCTTCAGGGTCGGCACCTAATACAAAAGGAAAATTTGTCAGGTTTCCTTCCGTAGTTTCATTCACTATTATTGAATCC
 -246 TTTGGCTACCATTCTTGAGAAACACAAACACTTCTTATATCTGTTCTACACAATTCTCTGAGTGCGTGCCACAGTTTGGTATCTTCATGATTGCTCATTGTTCATGCCCA
 -136 TAAGGAACATGTAACTTCCTCATTTATTTATTATTGCTTTTGTTTTCTTCTCACTAGTTAACTTTCGTTTCCCTATATAAACCCTCCTTTGTTCCCTTCCCTTCCCATCT
  -26 TCCATTTATTGATTCCAAACACAAACATGGCAATGAAGAAGCTCTTGTGGGTTGTTCTGTCCCTTTCTTTGGTTCTTGGAGTGGCCAACAGCTTTGATTTTCATGAGAAG
                    *          M  A  M  K  K  L  L  W  V  V  L  S  L  S  L  V  L  G  V  A  N  S  F  D  F  H  E  K
   85 GATTTGGAGTCCGAGGAAAGCTTGTGGGACTTGTACGAGAGATGGAGGAGTCACCACACGGTTTCGCGAAGCCTCGGTGAGAAGCACAAACGGTTTAACGTGTTCAAAGC
        D  L  E  S  E  E  S  L  W  D  L  Y  E  R  W  R  S  H  H  T  V  S  R  S  L  G  E  K  H  K  R  F  N  V  F  K  A
  195 AAATGTTATGCATGTGCATAACACTAACAAGATGGATAAGCCTTATAAGCTGAAACTGAACAAGTTCGCTGACATGACCAATCATGAATTCAGGAGTACCTATGCAGGCT
        N  V  M  H  V  H  N  T  N  K  M  D  K  P  Y  K  L  K  L  N  K  F  A  D  M  T  N  H  E  F  R  S  T  Y  A  G
  305 CAAAGGTTAATCACCACAAAATGTTCCGAGGCTCGCAACATGGGAGCGGTACCTTCATGTATGAGAAGGTTGGGAGTGTTCCTGCTTCAGTGGATTGGAGGAAGAAAGGT
        S  K  V  N  H  H  K  M  F  R  G  S  Q  H  G  S  G  T  F  M  Y  E  K  V  G  S  V  P  A  S  V  D  W  R  K  K  G
  415 GCTGTCACTGATGTGAAAGATCAGGGCCAATGTGgtaagcttaatgatcaccactcattaatcatagagctcaaactagctgaataatgtctgataattagctttttaatc
        A  V  T  D  V  K  D  Q  G  Q  C
  525 gaatttaatggcactgtatgatctatatagttgaattcatctaatgggataaaactgttgttattgagtgaaagctgcataaggtgggttttgtgtaatcatattgtttg

  635 gtactctcaaagatatcgttcagtctactacggcgctataaattgatttgatgtcctcacataggctatagtttctaataacattgttttctttgtcttgtgtagGTAGC
                                                                                                              G  S
  745 TGCTGGGCGTTTTCAACTATAGTAGCTGTTGAAGGCATTAACCAAATCAAGACAAATAAGTTGGTGTCATTGTCTGAACAAGAGTTGGTCGATTGTGACAAAGAAGAAAA
        C  W  A  F  S  T  I  V  A  V  E  G  I  N  Q  I  K  T  N  K  L  V  S  L  S  E  Q  E  L  V  D  C  D  K  E  E  N
  855 CCAAGGATGTAATGGTGGGTTGATGGAATCTGCTTTTGAGTTCATCAAACAGAAGGGAGGCATAACAACAGAAAGCAATTATCCATACACAGCACAAGAAGGAACGTGTG
        Q  G  C  N  G  G  L  M  E  S  A  F  E  F  I  K  Q  K  G  G  I  T  T  E  S  N  Y  P  Y  T  A  Q  E  G  T  C
  965 ACGAATCAAAGgtgcatgttaggccaaacatattttccaatgaaaccactatatttaatagatatacttgagaaactaaagtgccaaaaatctttcatgtggtagGTGAAT
        D  E  S  K                                                                                                V  N
 1075 GACCTAGCTGTGTCAATTGATGGTCATGAAAATGTCCCTGTGAACGATGAAAATGCACTGCTCAAAGCTGTCGCCAACCAACCTGTTTCTGTAGCCATTGATGCTGGGGG
        D  L  A  V  S  I  D  G  H  E  N  V  P  V  N  D  E  N  A  L  L  K  A  V  A  N  Q  P  V  S  V  A  I  D  A  G  G
 1185 GTCTGATTTCCAGTTCTACTCTGAGgtaaatatctaatatgcttataaagtgtaatttttttcgaaatttgttggaagaaaagcatccgttaccacttaacatgttacta
        S  D  F  Q  F  Y  S  E
 1295 taagtaatttaactattaggtgtcaatactgaaactatcctaatttaatataaaaaattttgcagGGAGTGTTTACGGGTGACTGTAACACAGATCTAAATCATGGTGTA
                                                                                                  G  V  F  T  G  D  C  N  T  D  L  N  H  G  V
 1405 GCAATTGTGGGATATGGAACAACGGTTGATGGAACTAATTATTGGATAGTGAGAAACTCTTGGGGACCAGAATGGGGAGAACAGGGTTACATCAGAATGCAAAGGAACAT
        A  I  V  G  Y  G  T  T  V  D  G  T  N  Y  W  I  V  R  N  S  W  G  P  E  W  G  E  Q  G  Y  I  R  M  Q  R  N  I
 1515 ATCTAAAAAGGAGGGACTTTGTGGCATAGCCATGATGGCTTCCTATCCAATCAAAAATTCCTCTGACAATCCAACAGGATCTTTGTCATCTCCCAAAGATGAACTTTGAA
        S  K  K  E  G  L  C  G  I  A  M  M  A  S  Y  P  I  K  N  S  S  D  N  P  T  G  S  L  S  S  P  K  D  E  L
 1625 AGCCTTCTACAAAATCAAACTCTCATAAGTCTTAGTTGTTGTTTTTTAACTTTTATCTAAGATTGAATGCGAATGTATCCATCAATATAATACATAACATACAAATAAGT
 1735 GTTTTAATTTGATTTGAGTGCTAGCAGTTTCACATGTACCCTTCCTTATAGCAAAGGGTTGAATGCAAAGATATAGCTTTAAACCATTAATAAAAAAAGTATATGATATG
                                                                                                                    *
 1845 ATCACTTTGAAGATACACAATGACATTTTATTTTTGGATAGACTATTATATTGGAACACGCATTAGATGACAGATAGAAGCTAGACATGATGAAATATGTAGCAAGAAAA
 1955 TTATCAGAATTGATCCATGATCAGCCATAGCACCTGATGGGATCAGTGCAACCTTTATCTGCTGACTGTTTGTTGCTTTCACCTTTTCCTCCACCCTTCTCCAGATATTG
 2065 CTGATGGTAGTCCTCTGCTTTGTAAAACCTCTTAGCTGGAAGAATTTCTGTCACAATCTTCTCCTTCAACTCCAACTGTTTTGCTTCTTTCGATTCCTGGGCTAAACGAG
 2175 CCTGCTCTTCGTCGTAGTAGTATATTCCTGATCTATACTGTGCACCCACATCATTGCCCTGAGATTTATTCATACAGAATAAACAACCAAAAATATAACATAATATTATC
 2285 ATAAATAATATAATTAAAAAGAAAAAAATAAAAACTTAAAATCCATTAGCCGCATGTTCATATTGGCGTAATTTACGTACCTGGCGGTTGATGGAAGTTGGGTCATGTCG
 2395 AGACCAAAACAGAGCAAGCAGATCCGTGTAGGGGGAAACTTCGGGTCGAACTGAACCCGAACCACTTCAACGTGGTTAGTGGTTCCGGTGCAAACCAGTTTGTAATCTGG
 2505 ATCAGGTGTGTGGCCCTGAGTATATCCGACCTCGGTCTTTACGACTCCCAGAATTC
```