Sequence of a cDNA encoding the platelet aggregation inhibitor trigramin

Michael P.Neeper and Marlene A.Jacobson*
Department of Pharmacology, Merck Sharp and Dohme Research Laboratories, West Point, PA 19486, USA

Submitted June 13, 1990 EMBL accession no. X51530

The cDNA for trigramin was cloned from a λZAP library prepared from polyA$^+$ or total mRNA extracted from the venom glands of *Trimeresurus gramineus*. The 2017 nucleotide (nt) insert has an open reading frame of 1053 nucleotides (nt) containing the coding region for the 72 aa mature trigramin determined from venom purified protein (1). The sequence contains seven additional amino-terminal residues which share homology with the platelet aggregation inhibitor, bitistatin isolated from the venom of *Bitis arietans* (2). An additional C-terminal alanine residue was identified in the cDNA clone which is not observed in venom purified trigramin. The putative signal sequence (underlined) is 1146 nucleotides upstream from the mature coding region and suggests that trigramin may be translated as a precursor protein which is processed post-translationally. The 3′ untranslated region contains a polyadenylation signal (AATAAA) 16 nucleotides upstream from the poly-A region.

REFERENCES

1. Huang,T.-F. *et al.* (1989) *Biochemistry* **28**, 661–666.
2. Shebuski,R.J. *et al.* (1989) *J. Biol. Chem.* **264**, 21550–21556.

```
   1  CGCTCAGGTTGGCTTGAAAGCAGGAAGAGATTGTCTGTCTTCCAGCCAAATCCAGTCTCC   60
  61  AAAATGATCCAAGTTCTTTTGATAACCATATGCTTAGCAGTTTTTCCTTATCAAGGGAGC  120
        M  I  Q  V  L  L  I  T  I  C  L  A  V  F  P  Y  Q  G
 121  TCCATAATCCTGGAATCTGGGAATCTTAATGATTATGAAGTAGTGTATCCAGAAAAAGTC  180
 181  ACTGCATTGCCCAAAGGAGCTGTTCAGCAAAAGTATGAAGACGCCATGCAATATGAATTT  240
 241  AAGGTGAATGGAGAGCCAGTGGTCCTTCACCTGGAAAAAAATAAAGGACTTTTTTCAGAA  300
 301  GATTACAGCGAGATTCATTATTCCCCTGATGGCAGAGAAATTACAGCATACCCTTCGGTT  360
 361  GAGGATCACTGCTATTATCATGGACGCATCGAGAATGATGCTGACTCAACTGCAAGCATC  420
 421  AGTGCATGTGATGGTTTGAAAGGACATTTCAAGCTTCAAGGGGAGATGTACCTTATTGAA  480
 481  CCCTTGGAGCTTTCCGACAGTGAAGCCCATGCAGTCTTCAAATATGAAAATGTAGAAAAA  540
 541  GAGGACGAGCCCCCCAAAATGTGTGGAGTAACCCAGAATTGGGAATCATATGAGTCCACC  600
 601  AAAAAGGCCTCTCAGTTAAATGTTACTCCTGAACAACAAAGATTCCCCCAAAGATACATT  660
 661  AAGCTTGGTATATTTGTGGACCACGGAATGTACACAAAATACAGTGGCAATTCTGAAAGG  720
 721  ATAACAAAAAGGGTACATCAAATGATCAACAATATAAATATGATGTGCAGAGCTCTGAAT  780
 781  ATTGTTACAACACTGAGTGTCCTAGAAATTTGGTCCGAAAAAGATTTGATTACGGTGCAG  840
 841  GCATCAGCGCCTACTACTTTGACCTTATTTGGAGCCTGGAGAGAGACAGTCTTGCTGAAT  900
 901  CGCACCAGTCATGATCATGCTCAGTTACTCACGGCCACTATCTTCAATGGAAACGTTATA  960
 961  GGAAGGGCTCCCGTGGGCGGTATGTGTGACCCGAAGCGTTCTGTAGCAATTGTTCGGGAT 1020
1021  CATAACGCAATAGTTTTTGTGGTTGCAGTTACAATGACCCATGAGATGGGTCATAATCTG 1080
1081  GGCATGCATCATGATGAAGATAAATGTAATTGTAACACATGCATTATGTCTAAAGTGTTA 1140
1141  AGCCGCCAACCTTCCAAATATTTCAGCGAATGTAGTAAGGATTATTATCAGACATTTCTT 1200
1201  ACTAATCATAACCCACAATGCATTCTCAATGCACCCTTGAGAACAGATACTGTTTCAACT 1260
1261  CCAGTTTCTGGAAATGAACTTTTGGAGGCGGGAGAAGATTGTGACTGTGGCTCTCCTGCA 1320
        V  S  G  N  E  L  L  E  A  G  E  D  C  D  C  G  S  P  A
1321  AATCCGTGCTGCGATGCTGCAACCTGTAAACTGATACCCGGGGCGCAGTGTGGAGAAGGA 1380
      N  P  C  C  D  A  A  T  C  K  L  I  P  G  A  Q  C  G  E  G
1381  CTGTGTTGTGACCAGTGCAGCTTTATAGAAGAAGGAACAGTATGCCGGATAGCAAGGGGT 1440
      L  C  C  D  Q  C  S  F  I  E  E  G  T  V  C  R  I  A  R  G
1441  GATGACCTGGATGATTACTGCAATGGCAGATCTGCTGGCTGTCCCAGAAATCCCTTCCAT 1500
      D  D  L  D  D  Y  C  N  G  R  S  A  G  C  P  R  N  P  F  H
1501  GCCTAACCAACAATGGAGCTGGAATGGTCTGCAACAGCAACAGGCAGTGTGCTGATGTGA 1560
      A  *
1561  ATACAGCCTACTAATCAACCTCTGGCTTCTCTCAGATTTGATTTTGGAGATCCTCCTTCG 1620
1621  AGAAGGTTTGGCTTCCCTGTAGTCCAAAGAGACCCATCTGCCTGCATCCTACTAGTAAAT 1680
1681  CACTCTTAGCTTTCATATGGAATCTAAATTCTGCAATATTTCTTCTCCATATTTAATCTG 1740
1741  TTTACCTCTTGCTGTAATCAAGCCTTTTTCCCACCACAAAGCTCCATGGATATGTACAAC 1800
1801  ACCAAGGGCTTATTTGCTGTCAAGAAAAAAAAATAGCCATTTTACCGTTTCCCAATTCCA 1860
1861  GAGCACATTTAATGCAACAGGTTCTGCCTTTAGAGCTGGTGTATTCAAAGTCAATGTTTC 1920
1921  CTCTCCCTCTCCCAAAATTTCATGCTGGCTTTCCAAGATGTAATTGCTTCCATCAATAAA 1980
1981  CTCACTATTCTCATTCAAAAAAAAAAAAAAAAAAAA 2017
```

* To whom correspondence should be addressed