© 1990 Oxford University Press

Cloning and sequencing of a zinc finger cDNA expressed in mouse testis

Daniel Nelki, Keith Dudley, Philip Cunningham and Mehdi Akhavan
Division of Biomolecular Sciences, King's College, Campden Hill Road, Kensington, London W8 7AH, UK

Submitted May 4, 1990

EMBL accession no. X52533

From a mouse testis cDNA library we have isolated a cDNA clone coding for a zinc finger protein of the C2H2 type (1, 2). The polypeptide encoded by this clone has outside the zinc finger region an N-terminal portion of 73 amino acids of which 25% are basic. There are 12 zinc fingers in the open reading frame with 9 out of the 11 linkers having the sequence Thr-Gly-Glu-Lys-Pro, indicating it is a member of the Kruppel family. At the C-terminus there are 9 amino acids after the last finger of which 6 are acidic and none basic. The open reading frame is bounded by stop codons at nucleotide 12 (upstream from the initiator methionine) and at nucleotide 1253.

REFERENCES

1. Evans,R.M. and Hollenberg,S.M. (1988) *Cell* **52**, 1–3.
2. Berg,J.M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 99–102.

```
      TGCATCTGATGTGAACAAAGATGAAATTCCAACTAGAAAGAAATGCGACAAGTTACCCAACAATAAGTTGTCTGATAAAGGTGACAAAAA
1     ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 90
                            M  K  F  Q  L  E  R  N  A  T  S  Y  P  T  I  S  C  L  I  K  V  T  K  T

      CCAAACCAGCAAAAAATGTGAGAAAGTATGCCCGTCATAGTGCATCCCATACCAAGGAAGACAAAATTCAGACCGGGGAGAAACGGAAAT
91    ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 180
        K  P  A  K  N  V  R  K  Y  A  R  H  S  A  S  H  T  K  E  D  K  I  Q  T  G  E  K  R  K  S

      CACACTGCCGTACTCCATCTAAACCTGAAAAAGCCCCAGGTTCTGGGAAACCTTATGAATGTAACCACTGTGGGAAGGTCCTCAGCCATA
181   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 270
        H  C  R  T  P  S  K  P  E  K  A  P  G  S  G  K  P  Y  E  C  N  H  C  G  K  V  L  S  H  K

      AACAGGGACTCCTTGACCATCAAAGAACTCACACTGGGGAGAAACCATATGAATGTAATGAATGTGGGATAGCTTTCAGCCAGAAGTCCC
271   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 360
        Q  G  L  L  D  H  Q  R  T  H  T  G  E  K  P  Y  E  C  N  E  C  G  I  A  F  S  Q  K  S  H

      ACCTTGTTCTACATCAGAGAACTCACACTGGGGAAAAACCATACGAGTGTGAACAGTGTGGCAAAGCACACGGACATAAACATGCCCTCA
361   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 450
        L  V  L  H  Q  R  T  H  T  G  E  K  P  Y  E  C  E  Q  C  G  K  A  H  G  H  K  H  A  L  T

      CTGACCATCTAAGAATCCATACTGGAGAAAAGCCCTACAAATGTAATGAATGTGGCAAAACGTTTAGACACAGCTCAAACCTTATGCAAC
451   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 540
        D  H  L  K  I  H  T  G  E  K  P  Y  K  C  N  E  C  G  K  T  F  R  H  S  S  N  L  M  Q  H

      ACCTAAGATCTCACACGGGTGAGAAGCCGTATGAATGTAAGGAATGTGGCAAATCCTTTAGATATAATTCATCTCTTACTGAACATGTGA
541   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 630
        L  R  S  H  T  G  E  K  P  Y  E  C  K  E  C  G  K  S  F  R  Y  N  S  S  L  T  E  H  V  R

      GAACACACACAGGTGAAATACCATACGAATGTAACGAATGTGGCAAAGCTTTCAAGTATGGCTCATCCCTGACTAAACATATGCGGATTC
631   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 720
        T  H  T  G  E  I  P  Y  E  C  N  E  C  G  K  A  F  K  Y  G  S  S  L  T  K  H  M  R  I  H

      ATACAGGGGAGAAACCATTTGAATGTAATGAATGTGGGAAAACTTTTAGCAAAAAGTCACACCTAGTTATACATCAAAGAACTCATACAA
721   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 810
        T  G  E  K  P  F  E  C  N  E  C  G  K  T  F  S  K  K  S  H  L  V  I  H  Q  R  T  H  T  K

      AGGAGAAACCTTATAAATGTGATGAGTGTGGGAAAGCCTTTGGACATAGCTCATCTCTTACCTACCATATGAGAACTCATACAGGTGACT
811   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 900
        E  K  P  Y  K  C  D  E  C  G  K  A  F  G  H  S  S  S  L  T  Y  H  M  R  T  H  T  G  D  C

      GCCCCTTTGAATGTAATCAATGTGGTAAGGCCTTTAAACAGATTGAAGGCCTTACCCAACACCAGAGAGTTCACACAGGGGAGAAACCCT
901   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 990
        P  F  E  C  N  Q  C  G  K  A  F  K  Q  I  E  G  L  T  Q  H  Q  R  V  H  T  G  E  K  P  Y

      ATGAGTGTGTTGAATGTGGGAAAGCCTTTAGTCAGAAGTCACACCTCATCGTACACCAGAGAACTCATACAGGGGAGAAACCCTTTGAAT
991   ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1080
        E  C  V  E  C  G  K  A  F  S  Q  K  S  H  L  I  V  H  Q  R  T  H  T  G  E  K  P  F  E  C

      GTTATGAGTGTGGAAAAGCCTTCAATGCAAAATCACAACTTGTTATTCATCAGAGATCCCACACTGGAGAGAAACCCTATGAATGTATTG
1081  ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1170
        Y  E  C  G  K  A  F  N  A  K  S  Q  L  V  I  H  Q  R  S  H  T  G  E  K  P  Y  E  C  I  E

      AATGTGGTAAAGCCGTCAAGCAAAATGCCTCTCTTACCAAACATATGAAAATTCACTCAGAAGAACAATCTGAGGAAGAAGATTAATGTA
1171  ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1260
        C  G  K  A  V  K  Q  N  A  S  L  T  K  H  M  K  I  H  S  E  E  Q  S  E  E  E  D  *

      GGAAACCTGACGAACTGACTGGTTTGGTATTATTTAACCTTAAAGATGTTCTCAATTTGATGATGTTAGAATATCTTTTTTTAGGAAATC
1261  ---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1350

      ATTCCTGGTGATACATGAGAGAATTTGAATATGGATCTTTACATAAGATGGTAATAAAATTTAACCTTGATCCCAAAAAAAAAA
1351  ---------+---------+---------+---------+---------+---------+---------+---------+---- 1434
```