

# Nucleotide and amino acid sequences of embryonic rat MAP2c

Thierry Doll, Anastasia Papandrikopoulou and Andrew Matus  
 Friedrich Miescher-Institut, PO Box 2543, 4002 Basel, Switzerland

Submitted December 15, 1989

EMBL accession no. X17682

Embryonic MAP2c, the low molecular weight form of microtubule associated protein MAP2, is encoded by a 6 kb mRNA transcript (1). The sequence of 1644 nucleotides including the entire coding region (1401 bp) of the MAP2c are shown below. The open reading frame starting at the first 5' ATG encodes a 467 residue protein of a predicted Mr=49,299 Da. Sequences were obtained from a cDNA clone, 38a (2) and a 700 bp fragment recovered from C6 glioma cell mRNA by the polymerase chain reaction (PCR) (3). The 700 bp fragment contained 5' non coding sequences and extended 3' past the alternate splice-junction in MAP2c (1). Comparison of the rat MAP2c nucleotide and amino acid sequence to cDNA clones encoding the high molecular

weight rat MAP2 shows a 100% identity. The homology between the mouse (4) and the rat nucleotide sequences is approximately 94%. For more details see ref. 1. The MAP2c splice-site is marked by an arrow; the three repeats encoding the microtubule binding-site motif (4) are underlined.

## REFERENCES

1. Papandrikopoulou, A., Doll, T., Tucker, R.P., Garner, C.C. and Matus, A. (1989) *Nature* **340**, 650-652.
2. Garner, C.C. and Matus, A. (1988) *J. Cell Biol.* **106**, 779-783.
3. Frohman, M.A., Dush, M.K and Martin, G.R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.
4. Wang, D., Lewis, S.A. and Cowan, N.J. (1988) *Nucl. Acids Res.* **16**, 11369.

```

1  GAATTCACATAACAACCGTGCATTACTTTACAACCTTGATTAGGAGACAGTACAGAGATCTGAAGGATGCTGACGAGAGGAAAGCAGGAAGGAAAGGCAC
      * M A D E R K D E G K A P
101  CACACTGGACATCAGCCTCACTCACAGAGGCAGCTGCACACCCCACTCGCCAGAGATGAAGGACCAGGGAGGCTCAGGGGAAGGGCTGAGCCGACGCGC
      H W T S A S L T E A A A H P H S P E M K D Q G G S G E G L S R S A
201  CAATGGATTTCATACAGAGAGAGGAGGAAGGCGCCTTTGGGAGCAGCGGTCACAGGACCTATTTCAGATACCAAGAGAACGGATCAACGGAGAG
      N G F P Y R E E E E G A F G E H G S Q G T Y S D T K E N G I N G E
301  CTGACCTCAGCTGACAGAGAAACAGCAGAGGAGTGTCTGCAAGGATAGTTCAAGTAGTCACAGCTGAAGCTGTAGCAGTCTGAAAGGTGAACAAGAGA
      L T S A D R E T A E E V S A R I V Q V V T A E A V A V L K G E Q E K
401  AGGAGGCCCAACACAAGGATCAGCCTGCAGCTCTGCCTTTAGCAGCTGAAGAAACAGTTAATCTGCCACCTTCCCCACCACCATCGCCAGCATCAGAACA
      E A Q H K D Q P A A L P L A A E E T V N L P P S P P P S P A S E Q
501  AACAGCTGCATGGAAGAAGCAACAAGTGGTGAATCAGCTCAGGCTCCAGTGCCTTAAACAGGCGAAGGACAAAGTCACTGATGGAATAACCAAGAGC
      T A A L E E A T S G E S A Q A P S A F K Q A K D K V T D G I T K S
601  CCAGAAAACGTTCTTCCCTCCCAAGACCTTCTCCATCCTCCCTCCTCGCAGGGCGGTATCAGGAGACAGGGAGGAGAACTCGTTCTCTGAAACAGCT
      P E K R S S L P R P S S I L P R R G V S G D R E E N S F S L N S S
701  CCATCTTTCAGCAGCGGACCACAGCTCAGAACAATTCGACAGCAGGAAAAGCGGCACCTCAACACCTACTACCCTGGATCTACTGCAATCAC
      I S S A R R T T R S E P I R R A G K S G T S T P T T P G S T A I T
801  CCCTGGCACTCCTCCTGCTACTCTTCACGTACCCAGGACCCCTGGAACCCCGAGCTATCCAGGACACCAGGAACCCCAATCTGGCATCTTGGTG
      P G T P P S Y S S R T P G T P G T P S Y P R T P G T P K S G I L V
901  CCCAGTGAGAAGAAAGTTGCCATCATTGCACTCCTCCAAAGTCCCAAGCTACTCCCAAGCAGCTTCGGCTCATTAAACCACTCTGCCAGACCTGAAGA
      P S E K K V A I I R T P P K S P A T P K Q L R L I N Q P L P D L K N
1001  ACGTCAAGTCCAAATCGGATCAACCGACAACATCAAAATACCAGCCTAAGGGGGTTCAGGTACAATTTGTTACTAAGAAGATAGACTTAAGCCATGTGAC
      V K S K I G S T D N I K Y Q P K G G Q V Q I V T K K I D L S E V T
1101  TTCCAAATGTGGCTCTCTAAGAACATCCGTCACAGGCCAGGTGGTGGACGCGTGAAGATTGAGAGTGTAAAGCTGGATTTCAAGGAGAAGGCCCAAGCT
      S K C G S L K N I R E R P G G G R V K I E S V K L D F K E K A Q A
1201  AAAGTGGCTCACTTGACAATGCTCACCATGTACTGGAGTGGTAAACGTTGACAGCCAAAGCTGAACTCCGAGAGCATGCAAGGCCCGCGG
      K V G S L D N A H H V P G G G N V K I D S Q K L N F R E H A K A R V
1301  TCGACCAGGGGCTGAGATCATCACAGTCGCCAAGCAGGTCAAGCGTGGCGTCTCCCGGCGACTCAGCAATGTCTCCTCTTGGAAAGCATCAACCT
      D H G A E I I T Q S P S R S S V A S P R R L S N V S S S G S I N L
1401  GCTCGAATCCCTCAGCTGGCCACTTTGGCTGAGGACGTCAGTCCGGCGCTCGCTAAGCAGGGCTTGGATCCTTCTCATCTAGCACTGAAAGTACCATT
      L E S P Q L A T L A E D V T A A L A K Q G L *
1501  TAGGCATGAGCTCTTGGCAGGAGTGGGCTCTGAGCAGGTGTATATTCTTTTACAACCATAGATAAATAATCTCATCCCCCACTGTAGTAAATTGT
1601  TACAATTTTATATAAAAAAAGAAATAGTATATGCAGAAAATG
    
```