The EMBO Journal vol.7 no.5 pp.1345 - 1353, 1988

The primary structure of the murine multifinger gene *mKr2* and its specific expression in developing and adult neurons

Kamal Chowdhury, Gregory Dressler, Georg Breier, Urban Deutsch and Peter Gruss

Department of Molecular Cell Biology, Max Planck Institute of Biophysical Chemistry, Am Fassberg, D-3400 Göttingen

Communicated by J.Tooze

The complete amino acid sequence of the murine finger-containing gene *mKr2* was determined. On the basis of sequence similarities in the repeated finger domain, *mKr2* belongs to the same class of developmentally expressed genes as *Drosophila Krüppel* and *hunchback*. The presence of metal ion and DNA-binding finger domains similar to those identified in TFIIIA supports the hypothesis that these genes regulate transcription. *mKr2* transcripts are restricted to neurons in the central and peripheral nervous system of adult animals. Furthermore, *mKr2* transcripts can be detected in all the major structures of the developing nervous system during embryogenesis. The data are consistent with the hypothesis that *mKr2* is a regulatory factor required for the differentiation and/or phenotypic maintenance of neurons.

Key words: *mKr2*/multifinger gene/cDNA/differentiation/neuron

Introduction

Given the complexities inherent in control systems required for the correct development of multicellular organisms, it is improbable that such mechanisms would evolve separately and independently of each other, and thus reasonable to assume that certain regulatory proteins and pathways may be conserved through evolution even among highly divergent species. In fact, conserved protein domains have been identified in a number of important regulatory genes including oncogenes, growth factors and receptors, and homeobox-containing genes (reviewed by Kahn and Graf, 1986; Gehring, 1985; Gehring and Hiromi, 1986; and references therein). As a result, it has been possible to isolate mouse (McGinnis *et al.*, 1984; Colberg-Poley *et al.*, 1985; Joyner *et al.*, 1985) and human (Levine *et al.*, 1984; Boncinelli *et al.*, 1985) genes using the conserved domains of the *Drosophila* homeotic genes *Antennapedia, Ultrabithorax* and *engrailed* as probes. Conversely, it has also been possible to isolate homologous genes from *Drosophila* using mouse genes as probes (Rijsewick *et al.*, 1987).

Recently, two members of a new gene family containing 'finger'-like nucleic acid-binding motifs have been isolated from a mouse genomic library using the *Drosophila Krüppel (Kr)* cDNA as a probe (Chowdhury *et al.*, 1987). The segmentation gene *Kr* belongs to the gap class, whose other members include *hunchback (hb)* and *knirps (kni)* (Nüsslein-Vollhard and Wieschhaus, 1980). Mutation at the *hb, Kr* and *kni* loci results in deletion of adjacent body segments in the anterior, middle and posterior embryonic regions respectively. Both *Kr* and *hb* contain finger-like domains (Rosenberg *et al.*, 1986; Tautz *et al.*, 1987) analogous to TFIIIA, a gene regulating the expression of 5S RNA in *Xenopus* (Ginsberg *et al.*, 1984). The finger domain of TFIIIA folds into nine finger-like structures that can bind DNA, RNA and Zn^{2+} (Hanas *et al.*, 1983; Brown *et al.*, 1985; Miller *et al.*, 1985; Frankel *et al.*, 1987). The *Kr* gene product is also a nuclear DNA-binding protein (Schuh *et al.*, 1986; Gaul *et al.*, 1987; Ollo and Maniatis, 1987), supporting the hypothesis that *Kr* has a transcriptional regulatory function.

In order to determine if the finger genes isolated from the mouse are, by analogy to *Drosophila*, controlling factors for morphogenesis and embryogenesis, it is first necessary to obtain a spatial and temporal expression pattern during development. Thus, an 8.5 day embryonic mouse cDNA library was screened with a mouse finger gene probe and the two resulting clones, cmKr2A and B, were subjected to further analysis. The complete amino acid sequence of *mKr2* was determined and its tissue-specific expression pattern in adults and developing embryos examined. Northern blot analysis of adult tissues demonstrates that the gene is expressed predominantly in the central nervous system (CNS); and *in situ* hybridization reveals that this expression is restricted to neuronal cell types. In addition, *mKr2* is expressed at high levels in the neural tube, mesencephalon, myelencephalon, diencephalon and spinal ganglia of developing embryos. These data are consistent with the hypothesis that the mouse finger gene, *mKr2*, plays an active role in the developmental regulation of the nervous system and/or the differentiation of neurons from embryonic ectoderm.

Results

Isolation and sequence of mKr2 cDNA clones

The mouse genomic clone, λmKr2, containing the finger domain, hybridizes to a 2.8 kb mRNA from undifferentiated embryonal carcinoma cells F9 (Chowdhury *et al.*, 1987) and mouse embryos. To obtain the *mKr2* cDNA clone, an 8.5 day p.c. mouse embryo mRNA library (kindly provided by K.Fahrner and B.Hogan) was screened with a cloned 362 bp *Eco*RI fragment from λmKr2, containing both unique and finger-coding sequences. The complete DNA sequence of two overlapping clones, cmKr2A and cmKr2B, is shown in Figure 1. Together, these two clones span 2.7 kb of the ~2.8 kb mRNA detected by Northern blotting.

The sequence presented in Figure 1 contains the entire protein-coding region of the *mKr2* gene. The first ATG, at nt 433, is a consensus translational start site, with the sequence CANNATGA/G (Kozak, 1983; Lütcke *et al.*, 1987),

```
1       GAATTCATGT CTTACGGTCA AGGCAGGAGT TACTGACATA CAAAGATGTG TCTGAGGACT TCACTTGGGA GGAGTCAATG
81      CAGCTATATT CTGCACAGAG GAACCTGAAT GGCAGAGTGA TGCTGGAGAA CTGTAGCACC TTGGTCTCCT TGGGTGAGCA
161     CCCCCGCCTT CAGCCATCTG CCATATGGAA TACACAGGGC ATCAGCCGTC CAAACCAAAT GTGATCTTGG AGCTGGGACA
241     AGAACTGACC CTGAAAAAAG GAGAATTCCC AGGTTGCATC TCTCTGGGCC TGGCTCGTTG TCTGTGAGCA TCCTGTGCTC
321     CCAAGAGGAT CTCCTAGACT GGGTGTTAAA ACTTTGCTGC GAGACAAAAC TTGAAACAAG AGAGTCAACT CCAAAGTCCA
                  .....TT TGTTTCCTTT TTCAGGCTGC GAGACAAAA. ...    genomic sequence
                                        splice

401     CCATTACTGG AATTTTGCCC TATGGGGTCA TA ATG GAC AGG GAA GAT CTC TGG CAT TCA GCT TTG GGG GCT
1                                           Met Asp Arg Glu Asp Leu Trp His Ser Ala Leu Gly Ala

472     GTC TGG GAT CCT ACC TGT TGG TTA AAG GGG CAA CAG GAG AGA TAC CTG GGC CAA GTG ACA GTG GCC
14      Val Trp Asp Pro Thr Cys Trp Leu Lys Gly Gln Gln Glu Arg Tyr Leu Gly Gln Val Thr Val Ala

538     CAA AAG GAA ATC TAT AAT GAG AAG AGT GTG TGT GGA GGT AAT ACA ACT GAG AAC TCT TCC ACT GAG
36      Gln Lys Glu Ile Tyr Asn Glu Lys Ser Val Cys Gly Gly Asn Thr Thr Glu Asn Ser Ser Thr Glu
                                                                          CHO

604     GGA TCA ATG CTT AAT ACA CCA CAA AGC ATT CCT GTG ACA CCT TGT AAT TGG AAT TCA TAT AGA AAA
58      Gly Ser Met Leu Asn Thr Pro Gln Ser Ile Pro Val Thr Pro Cys Asn Trp Asn Ser Tyr Arg Lys

670     GAT TCT AAA CAA AAT TCT GAG TTA ATG AAA ACT TCA AGA ATG TTT GTA CAA AAG AAA GTA TAT GGC
80      Asp Ser Lys Gln Asn Ser Glu Leu Met Lys Thr Ser Arg Met Phe Val Gln Lys Lys Val Tyr Gly

736     TGT GAT GAG TGT GGG AAA ACC TTC AGG CAG AGT TCA TCC CTG CTT AAG CAT CAG AGG ATT CAC ACT
102     Cys Asp Glu Cys Gly Lys Thr Phe Arg Gln Ser Ser Ser Leu Leu Lys His Gln Arg Ile His Thr

802     GGG GAG AAA CCC|TAC ACA TGC AAC GTG TGT GAC AAG CAT TTC ATC GAA CGC TCC TCC CTC ACT GTG
124     Gly Glu Lys Pro|Tyr Thr Cys Asn Val Cys Asp Lys His Phe Ile Glu Arg Ser Ser Leu Thr Val

868     CAT CAG AGG ACT CAC ACG GGA GAA AAA CCC|TAC AAA TGT CAT GAG TGT GGG AAA GCC TTT AGT CAG
146     His Gln Arg Thr His Thr Gly Glu Lys Pro|Tyr Lys Cys His Glu Cys Gly Lys Ala Phe Ser Gln

934     AGT ATG AAC CTT ACT GTT CAT CAG AGA ACT CAC ACT GGA GAG AAG CCT|TAT CAG TGC AAA GAG TGT
168     Ser Met Asn Leu Thr Val His Gln Arg Thr His Thr Gly Glu Lys Pro|Tyr Gln Cys Lys Glu Cys

1000    GGG AAA GCA TTC CGC AAG AAT TCA TCC CTT ATT CAA CAC GAA AGG ATT CAT ACT GGA GAG AAG CCC|
190     Gly Lys Ala Phe Arg Lys Asn Ser Ser Leu Ile Gln His Glu Arg Ile His Thr Gly Glu Lys Pro|
                                CHO

1066    TAC AAA TGT CAT GAT TGT GGG AAA GCT TTC AGT AAG AAT TCT TCC CTT ACT CAG CAT CGC AGA ATT
212     Tyr Lys Cys His Asp Cys Gly Lys Ala Phe Ser Lys Asn Ser Ser Leu Thr Gln His Arg Arg Ile
                                                  CHO

1132    CAC ACC GGT GAG AAA CCT|TAT GAG TGT ATG ATC TGT GGA AAG CAT TTC ACT GGA AGA TCA TCC CTA
234     His Thr Gly Glu Lys Pro|Tyr Glu Cys Met Ile Cys Gly Lys His Phe Thr Gly Arg Ser Ser Leu

1198    ACT GTA CAC CAA GTT ATA CAC ACT GGA GAG AAG CCC|TAT GAG TGC ACT GAA TGT GGA AAG GCA TTC
256     Thr Val His Gln Val Ile His Thr Gly Glu Lys Pro|Tyr Glu Cys Thr Glu Cys Gly Lys Ala Phe

1264    AGC CAA AGC GCC TAT CTC ATT GAA CAT CGA AGG ATT CAT ACT GGT GAA AAA CCC|TAT GAA TGT GAT
278     Ser Gln Ser Ala Tyr Leu Ile Glu His Arg Arg Ile His Thr Gly Glu Lys Pro|Tyr Glu Cys Asp

1330    CAG TGT GGG AAA GCT TTC ATT AAG AAC TCA TCC CTT ATA GTG CAC CAG AGG ATA CAT ACA GGA GAG
300     Gln Cys Gly Lys Ala Phe Ile Lys Asn Ser Ser Leu Ile Val His Gln Arg Ile His Thr Gly Glu
                                     CHO

1396    AAA CCC|TAT CAG TGC AAT GAA TGT GGA AAG CCC TTC AGT AGG AGT ACA AAC CTT ACA AGA CAT CAG
322     Lys Pro|Tyr Gln Cys Asn Glu Cys Gly Lys Pro Phe Ser Arg Ser Thr Asn Leu Thr Arg His Gln

1462    AGG ACT CAT ACA|TGAGAAA TGCTCTGGTC AACTCCTTCA TTTTGTTTAA CCTTTTGATA CTAATCAGAT GTCACACTGA
344     Arg Thr His Thr|STP

1541    TGTAGATACT TAGTAAATAT AACTAATGTG GTAACTTTTT AGTTGAAGAA CAATATAGCA TAGCAGATAA TAAAATCATA
1621    GAGAGTAACC ATGTGAAAGT GGAGAACCCT TGATTCAGAT TTGAGTAGAC TATGACATAA ATAATGAAGC TGTATTTGGT
1701    CACACTGGGA AATCTGATTT GGGCATCACA AAATTCACAT TGAAAGGACA ATGTGATAGT GATTAACTGG ACAGCCAAAA
1781    AATCTGATTC ACCCAGTTCT AACCGCCCAT GAGCCCTTTA TCTAGTAATG AGAGATTTGG AATATACAAA CTCTGTGGAT
1861    CTATATTATC CAAAATGCCC AACACGAAGA TTCAGTGAAT TTTTTTTTGC TGTTGTATTT GCTATGGAGA TGCTTTTATT
1941    TCATAGTATT TTTGGGGTGT TTTAGGTTAA AGTTTTCTAC CTCTGACTCT TACTGTCCTA CAAATCCCTG AATTACTAAC
2021    TCTAAAATGT AAGCTAGTTG CTACTGGGTG GCAGATAAAT TGTGCATATT TTAATGCCAT GCATGCTGCT TTGAGTTCTC
2101    AGCTAATAGT CATTAAACTG TTGTTAAATA AATTTACTCT ATTTAGATGA CAATGTCACT TAAATATAAT CTCTAATTAT
2181    CTAGGAATCA AATACATATC TTTGTGTCTT ATTGGAGAGT CATAGAAATT TAAATACAGA TGAATTGAAC ACTAATTGTG
2261    ATGGACAGTT ATTATGTCAA CTTGATTGTC AACTAAAGTC ATCTAAGAAG AGGGAATCTC AATAAGAAAA TGCCTTTATA
2341    AGATGGGCCA TAGGACTACC TACAGGGCAT TTTCCTAATT AGTGATTGAT AGGGGAGGAT CCAGCCATTG TGGGTGGTGC
2421    CATCCCTTGG CTGATGGTCC TGAGTTCTAT ACAGAAAGCA GGCTGAGCAA GCAGCACTCC TCATGGCATC AGCTCCAGCC
2501    TCTAGGTTCC TGCCTTGTTT AAGTTCCTGT CCTGACTTTC TTTGGTGATG AAAAGTGATA TAGACATGTA AGCCAAATAA
2581    ACCTTTTCCT CCACGAAAAA AAAAAAAAAA AAAAA
```

and is in the same reading frame as the finger domain. Furthermore, the 5′ leader sequence from nt 1 to nt 433 contains three stop codons in the same reading frame as the finger domain and multiple stop codons out of frame. A 347 amino acid open reading frame, terminating at nt 1473, is followed by 1125 nt of 3′ untranslated sequence, containing nine stop codons in all three reading frames. Three potential AAUAAA consensus polyadenylation signals (Proudfoot and Brownlee, 1976) are located in the 3′ untranslated region, and a poly (A) tail is found 14 nt downstream of the most distal polyadenylation signal at nt 2580. The smaller clone cmKr2B does not contain a poly (A) tract at its 3′ end and may represent internal priming as shown previously with this particular cDNA library (Krumlauf *et al.*, 1987). The cDNA sequence is colinear with the genomic sequence from nt 355 to the 3′ end of the cDNA. Upstream of nt 355, the genomic sequence is different, indicating the presence of an intron. A consensus eukaryotic splice acceptor sequence is present at nt 355, as indicated in Figure 1. Thus, *mKr2* has at least one 5′ untranslated exon.

Conceptual translation of the 347 amino acid open reading frame, initiating at nt 433 and terminating at nt 1473, is also shown in Figure 1. The predicted protein has a calculated mol. wt. of 38 kd. There are four potentially strong glycosylation signals represented by NSS, one of them in the unique region and the rest in the finger domain. The unique N-terminal domain which consists of 99 amino acids is followed by a 248 amino acid repetitive finger domain. The potential metal and nucleic acid-binding finger domain contains eight complete fingers of 28 amino acids each and ends with a partial finger of 24 amino acids. The individual fingers of *mKr2* have the following consensus sequence:

YxCxxCGKAFxxxxxLxx**H**QRI**H**TGEKP.

Further analysis of the protein, using the algorithm of Garnier *et al.* (1978), failed to detect other DNA-binding motifs, such as the helix-turn-helix motif present in several yeast and mammalian homeobox proteins (Johnson and Herskowitz, 1985; Laughon and Scott, 1984). The data suggest that the entire amino acid sequence of the multiple finger containing gene, *mKr2*, is present in the overlapping cDNA clones isolated. The presence of the potential DNA binding finger domain and its implication in the regulation of morphogenic and ontogenic events warrants a detailed temporal and spatial expression analysis in adult mice and developing embryos.

Expression of mKr2 is tissue and cell-type specific

If the mouse finger protein *mKr2* has regulatory function, its expression may be restricted to certain tissues or cell types. To test this hypothesis, Northern blot analysis was used to examine poly $(A)^+$ RNAs from a variety of adult tissues for *mKr2* transcripts. The probe used was a 362 bp *Eco*RI fragment, which detects a single 2.8 kb mRNA in F9 cells and does not cross hybridize to other 'finger'-like genes under the conditions used (Chowdhury *et al.*, 1987). The data presented in Figure 2 indicate high levels of the 2.8 kb *mKr2* transcript in adult brain and spinal cord, as well as the kidney. All other tissues examined—including heart, spleen, lung, liver, testis, muscle, ovaries and pancreas—do not contain detectable levels of *mKr2* transcripts.

To determine if *mKr2* is transcribed in a cell-type-specific manner in the mouse CNS, *in situ* hybridization was performed with frozen frontal sections of adult brain tissue. Sections were hybridized with [^{35}S]-UTP-labeled RNA probes corresponding to the sense and anti-sense polarity of the 362 bp *mKr2 Eco*RI fragment. With the anti-sense probe, strong hybridization is detected in regions containing densely packed neurons, such as the dentate gyrus of the hippocampus (Figure 3a and a') and the pyriform cortex (Figure 3b and b'). As can be seen in field CA3 of the hippocampus, it is clear that grains are clustered over and around large pale staining nuclei characteristic of pyramidal neurons in this region (Figure 3c). Small dark staining nuclei, more characteristic of neuroglial cells, do not show significant levels of hybridization. Individual neurons containing transcripts that hybridize to the mKr2 anti-sense probe are also found in the cerebral cortex (Figure 3d) and brain stem area (Figure 3e). Again, grains are found predominantly over and around large pale nuclei, usually containing one or two dark staining nucleoli. Adjacent serial sections hybridized with the *mKr2* sense probe were entirely negative (Figure 3f) as were RNase pretreated sections hybridized with the *mKr2* anti-sense probe (data not shown). All CNS sections were stained for Nissl substance with cresyl violet acetate; however, neuronal cell bodies are not visible due to the RNase digestion step in the post-hybridization washing procedure. These data indicate that *mKr2* transcripts in the CNS are primarily restricted to neurons, although we cannot rule out low levels of expression in other cell types or a limited subset of neurons not expressing *mKr2*.

Expression of mKr2 is restricted to the developing nervous system during embryogenesis

Expression of *mKr2* during murine embryogenesis was analyzed initially by Northern blotting of poly $(A)^+$ RNA from 8, 10, 12, 14 and 16 day p.c. embryos. RNA was extracted separately from embryonic and extra-embryonic tissue, and filters were hybridized with the 362 bp *Eco*RI fragment of *mKr2*. The data presented in Figure 4 show detectable levels of the 2.8 kb *mKr2* mRNA in 10 and 12 day embryos. To normalize for the amount of RNA per lane, the filter was rehybridized with a c-*ras*^{Ha} probe. Thus, the lower level of *mKr2* transcripts in the 10 day embryonic RNA is, in part, due to less RNA loaded. In poly $(A)^+$ RNA from 14 and 16 day p.c. embryos, *mKr2* transcripts could no longer be detected, as in all extra-embryonic tissue examined.

Hybridization of *mKr2* to frozen sections of mouse embryos also reveals a tissue-specific expression pattern during embryogenesis. Figure 5 presents *in situ* hybridization data obtained with 10, 12 and 17 day p.c. embryos using [^{35}S]-UTP-labeled sense and anti-sense RNAs of the 362 bp

Fig. 1. The nucleotide sequence of *mKr2* cDNAs and the predicted amino acid sequence of the protein. The nucleotide sequence is a composite of two cDNAs. cmKr2A starts at nt 346 and ends with a stretch of ~100 adenine residues at nt 2596. The clone cmKr2B starts at nt 1 and its 3′ end has not been determined. A splice site is located at nt 355. The genomic DNA sequence spanning the splice junction is shown below the cDNA sequence. Stop codons located in the protein reading frame and at the 5′ untranslated region are underlined. Potential glycosylation signals are marked by -CHO. and the polyadenylation signals are also underlined. The finger domain is bracketed and an individual finger is separated by a vertical line.

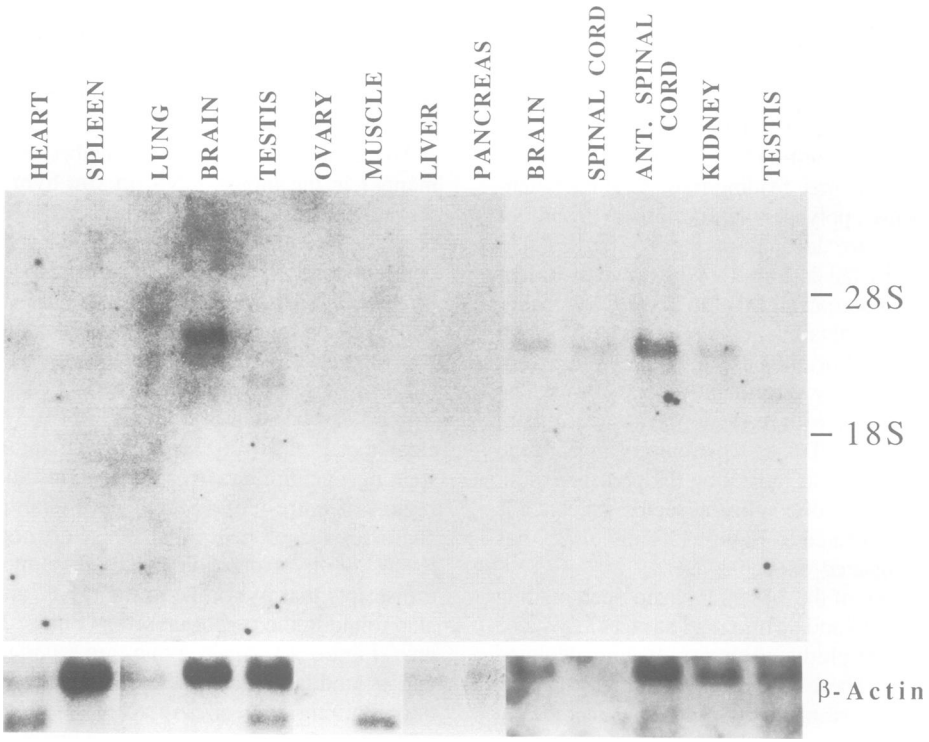


Fig. 2. Expression of *mKr2* sequences in adult mouse tissues as detected by Northern blotting. Poly $(A)^+$ RNAs were isolated from various adult mouse tissues as described in Materials and methods. Approximately 5 μg of each RNA was electrophoresed through an agarose–formaldehyde gel and blotted onto a GeneScreen Plus filter. The probe used for hybridization was the 362 bp *Eco*RI fragment (nt 655-1017, Figure 1).

*Eco*RI *mKr2* fragment. In the 10 day p.c. embryo, *mKr2* transcripts can be detected in the diencephalon (Figure 5a and a') as well as the neural tube. All of the major structures of the developing CNS at 12 days p.c. show high levels of *mKr2* transcripts (Figure 5b and b'). These include the diencephalon, mesencephalon, myelencephalon and neural tube as indicated. In addition, transcripts in the developing peripheral nervous system (PNS) also hybridize to *mKr2*, as seen in the thoracic spinal ganglia in parasagittal sections (Figure 5c and c'). Although Northern blot analysis shows a decrease in *mKr2* transcripts after day 12, they are detectable in 17 day p.c. embryos by *in situ* hybridization. Most, if not all, structures of the CNS and PNS contain *mKr2* transcripts at 17 days p.c. These include the cervical spinal ganglion (Figure 5d), the trigeminal ganglion (Figure 5e), the cerebral cortex, and the neural tube (data not shown). Particularly in the ganglia (Figure 5d and e) it is clear that *mKr2* transcripts are cell-type specific, as grains are found predominantly over and around large cell bodies with large pale nuclei, indicative of neuronal cells. The data indicate that cells of the developing C- and PNS accumulate *mKr2* transcripts beginning at ~10 days p.c. and continuing through development into adulthood. Thus, transcription of *mKr2* during development is restricted to cells originating from the ectoderm of the neural tube and the neural crest.

Discussion

The murine gene *mKr2* belongs to the class of multiple finger-containing genes based on structural similarities in the repetitive finger domain. Other members of this gene class include: the *Drosophila Kr, hb* and *Sry* (Vincent *et al.*, 1985) genes; the *Xenopus TFIIIA* and *Xfin* genes (Ruiz i Altaba *et al.*, 1987); and the yeast genes *GAL4* (Johnston, 1987), *PPRI* (Kammerer *et al.*, 1984), *q1F* (Baum *et al.*, 1987), *Lac 9* (Salmeron and Johnston, 1986), *ARGR11* (Messenguy *et al.*, 1986), *Leu3* (Friden and Schimmel, 1987) and *ADRI* (Hartshorne *et al.*, 1986). Sufficient available evidence suggests that individual finger-like structures bind one molecule of a metal ion and assume a tertiary structure capable of binding DNA (Miller *et al.*, 1985; Frankel *et al.*, 1987; Johnston, 1987). The nuclear localization and DNA-binding ability are indicative of transcriptional regulatory proteins, as demonstrated conclusively for *Kr* (Schuh *et al.*, 1986; Ollo and Maniatis, 1987). It is necessary to distinguish the multiple finger-containing genes from genes identified that contain one or two potential fingers, such as the human estrogen (Greene *et al.*, 1986), glucocorticoid (Green *et al.*, 1986), and mineralocorticoid receptors (Arriza *et al.*, 1987), and the c-*erb*A thyroid hormone receptor (Weinberger *et al.*, 1986; Sap *et al.*, 1987). Although it has been postulated that, on the basis of conserved cysteine and histidine residues, these receptor proteins bind DNA via a metal-dependent finger-like structure, no direct evidence for metal ion-mediated DNA binding exists.

The consensus finger sequence of *mKr2* is similar to that of *Drosophila Kr, hb* and *Sry*, *Xenopus xfin*, and yeast *ADRI*, but diverges from *Xenopus TFIIIA*. The H–C link sequence HTGEKPYXC (Schuh *et al.*, 1986), located between the individual finger domains, is conserved between *mKr2* and *Kr*, and thus facilitated the isolation of the genomic λmKr2 clone. Furthermore, antibodies generated against the H–C link sequence recognize nuclear proteins in a variety of organisms and cell types (Schuh *et al.*, 1986). This high degree of conservation suggests a function for the H–C link sequence, perhaps in maintaining the correct tertiary finger

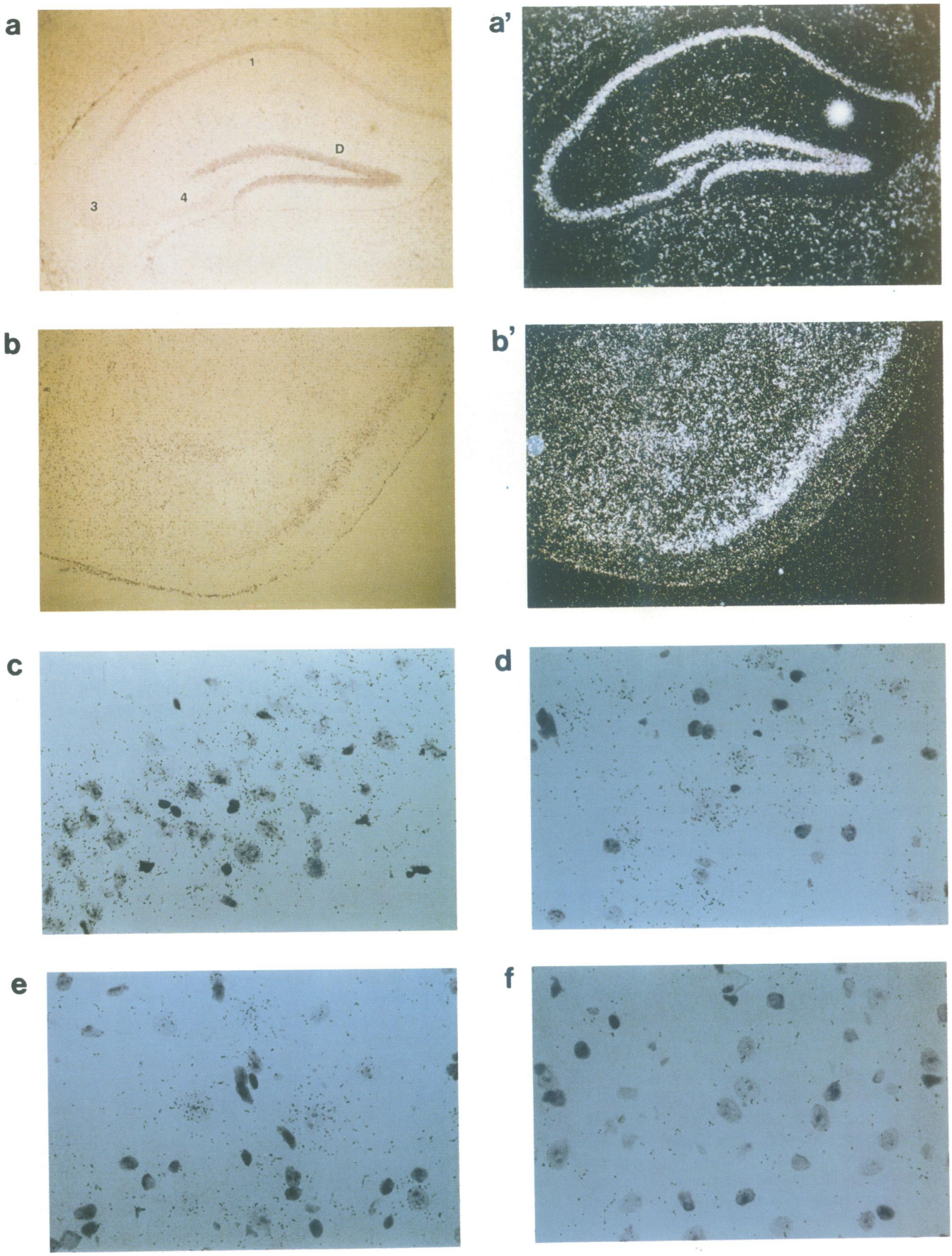


Fig. 3. *In situ* hybridization of *mKr2* in adult brain. Anti-sense probes were used for **a**–**e**; the sense probe is shown in **f**. (**a**) Frontal section showing hippocampus regions CA1 (1), CA3 (3), CA4 (4), and dentate gyrus (D); (**a'**) dark field image of **a**; (**b**) frontal section showing the pyriform cortex; (**b'**) dark field image of **b**; (**c**) high magnification of CA3 region in the hippocampus; (**d**) high magnification of neurons in the cerebral cortex; (**e and f**) high magnification of the brain stem area.

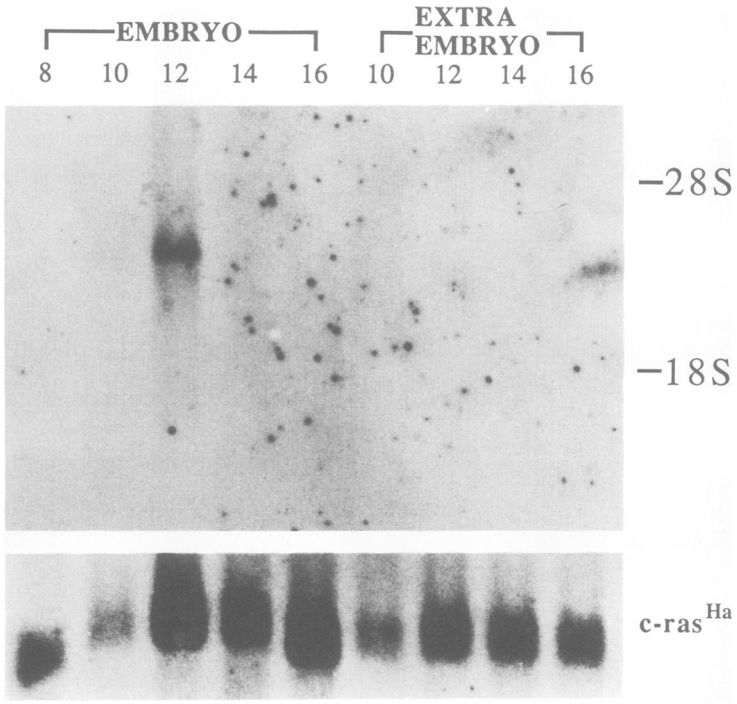


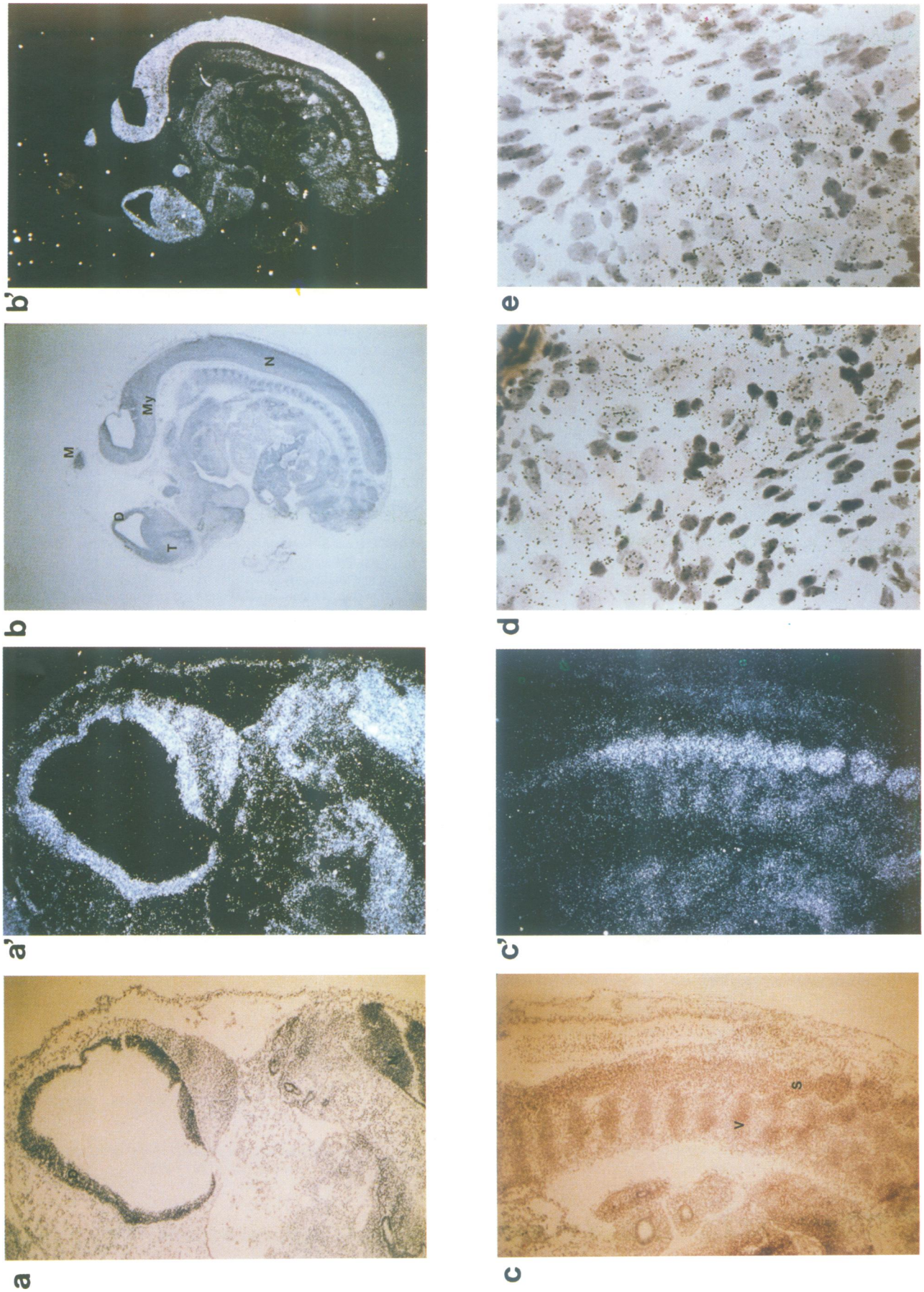
Fig. 4. Expression of *mKr2* sequences in developing mouse embryos as detected by Northern blotting. Poly $(A)^+$ RNA was isolated from 8 day old total embryo. The 10–16 day embryos were dissected into embryonic and extra-embryonic tissues and the poly $(A)^+$ RNAs were isolated from these tissues separately. Electrophoresis, blotting, and hybridization were done as described in the legend to Figure 2 and in Materials and methods. The lower panel shows a rehybridization of the same blot to a nick translated c-*ras*^{Ha} probe.

structure or in factor recognition. The finger domains of *mKr2, xfin, TFIIIA, Kr, hb, Sry beta* and *Sry delta* represent 71, 88, 76, 30, 25, 52 and 45% of the entire coding sequence, respectively. The positions of unique non-finger domains vary, as *mKr2* has a 99 amino acid N-terminal region, *Kr, TFIIIA*, and *xfin* contain both N- and C-terminal unique domains, and *hb* contains a unique domain inserted between two finger domains. A protein sequence comparison failed to reveal any significant similarities in the unique domains of the multiple finger-containing genes sequenced to date. The positions and number of unique domains may be indicative of different interactions these proteins establish with other cellular factors.

The murine gene *mKr2* is transcribed in neurons of the C- and PNS in adult animals. In addition, *mKr2* transcripts can be detected by *in situ* hybridization in the major structures of the developing nervous system beginning at 10 days p.c. The CNS is derived from the neural tube and the neural folds which are formed at 8 days p.c., whereas the spinal dorsal root ganglia and the cranial sensory ganglia are derived from the neural crest cells. Although *in situ* hybridization to 8.5 day p.c. embryos has proved inconclusive (data not shown), it is reasonable to assume that low levels of *mKr2* transcripts must be present since the clones cmKr2A and cmKr2B were isolated from an 8.5 day embryonic cDNA library. Northern blot analysis reveals a peak of *mKr2* transcripts at 12 days p.c. during embryogenesis. It is likely that this peak, and the subsequent decrease, is dependent on the ratio of developing nervous tissue to total body mass rather than a decrease in transcriptional activity, as 17 day p.c. embryos and adult animals still transcribe *mKr2* in neuronal cells. Neurons are formed from primordial cells of the neuroepithelium, or ependymal layer, that differentiate and migrate to the mantle layer of the developing CNS. It is generally believed that neuroblasts, the precursor of neurons, and neuroglial cells, which include astrocytes and oligodendrocytes, are derived from the same primordial ependymal cells. It would be interesting to examine precisely in what cell types *mKr2* expression can first be detected. If the neuronal precursor cells already express *mKr2*, then *mKr2* may be a necessary regulatory factor required for neuronal differentiation. Unfortunately, the cellular resolution for *in situ* hybridization to developing embryos is poor, in part due to the high cellular density in tissues undergoing rapid mitosis and the relatively low signals observed. A precise cell lineage analysis of *mKr2* expression will probably require a mono-specific antiserum against this protein.

A regulatory protein involved in morphogenesis and embryogenesis may have different functions depending on the time and place of its expression. The *Drosophila Kr* gene is expressed in the blastoderm of pregastrulation embryos, as well as in the developing nervous system after segmentation (Gaul *et al.*, 1987). Similarly, the *Drosophila* segmentation gene *fushi tarazu (ftz)* is expressed not only at the blastoderm stage but also later, during neurogenesis in the developing CNS (Carroll and Scott, 1985; Hiromi *et al.*, 1985). Three different regulatory elements, enhancer,

Fig. 5. *In situ* hybridization of *mKr2* anti-sense probes in developing embryos. (**a**) 10.5 day p.c. embryo showing diencephalon (D) and neural tube (N). (**a'**) Dark field image of **a**; (**b**) mid-sagittal section of a 12.5 day embryo showing neural tube (N), myelencephalon (My), mesencephalon (M), diencephalon (D), and telencephalon (T); (**b'**) dark field image of **b**; (**c**) parasagittal section of a 12.5 day p.c. embryo showing spinal ganglia (S) and pre-vertebrae (V); (**c'**) dark field image of **c**; (**d**) cervical spinal ganglion of a 17.5 day p.c. embryo; (**e**) trigerminal ganglion of a 17.5 day p.c. embryo.


b'
e
b
M
My
N
P
T
d
a'
c'
a
c
V
S

neurogenic element and zebra element, control the expression of *ftz* (Hiromi *et al.*, 1985). The expression of *mKr2* in the developing nervous system may signal the onset of neuronal differentiation. Alternatively, *mKr2* expression may be required for a terminally differentiated neuronal phenotype. The *mKr2* transcripts detected in the adult kidney may reflect transcripts originating from the adrenal medulla, a tissue derived from neural crest cells, since the adrenal gland was not removed from the tissue used for kidney RNA isolation. Interestingly, *mKr2* transcripts cannot be detected in the kidney or adrenal glands of 17 day p.c. embryos, suggesting that transcription increases after birth.

Low stringency hybridization to genomic mouse DNA has previously revealed the presence of multiple genes homologous to the finger domain of *mKr2* (Chowdhury *et al.*, 1987). As the first mammalian multifinger gene sequenced and characterized, *mKr2* is a member of a much larger gene family that may have evolved to regulate specific events in the complex processes of cellular differentiation and embryonic development. In *Drosophila*, the multifinger segmentation genes of the gap class, *Kr, hb* and *Kni*, interact with each other to regulate the spatial expression patterns during embryogenesis (Ingham *et al.*, 1986; Jäckle *et al.*, 1986). As more mammalian multifinger genes are isolated, sequenced and characterized in conjunction with functional assays currently under development, their potential role in the regulation of embryogenesis can be addressed.

Materials and methods

Isolation and sequencing of the cDNA clone cmKr2

An 8.5 day old C57BL whole mouse embryo cDNA library in λgt10 (Fahrner *et al.*, 1987) was screened with a 362 bp *Eco*RI DNA fragment from the 5′ end of the genomic *mKr2* clone (Chowdhury *et al.*, 1987). Approximately 4×10^5 phages were plated out and after transferring the plaques to nitrocellulose filters, the phage DNA was denatured and immobilized according to the standard procedures (Maniatis *et al.*, 1982). Hybridization to the nick-translated ^{32}P-labeled probe was carried out at 65°C in 3 × SSC, 5 × Denhardt's solution, 0.1% SDS, and 100 μg/ml of denatured salmon sperm DNA for 18 h. The filters were washed at 60°C in 2 × SSC and 0.1 × SSC, 0.1% SDS for 30 min each respectively, dried at 37°C, and exposed to Kodak X-Omat film at −70°C. Two positively hybridizing individual clones were isolated and plaque purified. Inserts from both these clones were subcloned in Bluescript plasmid vector (Stratagene) and sequenced by the method of Maxam and Gilbert (1977).

RNA isolation and Northern blotting

Tissues or embryos were dissected from NMRI mice, quickly frozen in liquid nitrogen and stored at −70°C. The RNA was isolated from the frozen tissue after disrupting it with a Polytron homogenizer using the guanidinium thiocyanate method (Chirgwin *et al.*, 1979). The homogenate was centrifuged through a 4 ml cushion of 5.7 M CsCl, 25 mM sodium acetate (pH 5.0) in a Beckman SW41 rotor at 31 000 r.p.m. for 22 h. The total RNA pellet was dissolved in 10 mM Tris, 1 mM EDTA and ethanol precipitated. Poly $(A)^+$ RNA was isolated by retention on oligo (dT) cellulose (Aviv and Leder, 1972). About 5 mg each of poly $(A)^+$ RNAs were loaded in each lane and electrophoresed through a 1% agarose gel in 3.7% formaldehyde and MOPS buffer (20 mM morpholine propane sulfonic acid, 50 mM sodium acetate, 10 mM EDTA; pH 7.0). The electrophoresis buffer was identical to the agarose gel buffer. After blotting the RNAs onto GeneScreen plus (Dupont-NEN), the membrane was hybridized at 42°C in 50% formamide, 1 M NaCl, 1% SDS, and 200 μg/ml denatured salmon sperm DNA, according to the manufacturer's protocol. In all cases, the same 362 bp *Eco*RI fragment consisting of both unique and finger domain sequences was used as the probe. The blots were washed twice in 2 × SSC at room temperature, twice in 2 × SSC, 1% SDS at 65°C for 30 min, and twice in 0.1 × SSC at 60°C for 30 days min. The filters were exposed to Kodak X-Omat films for 3−7 days.

In situ hybridization

The techniques employed for *in situ* hybridization were essentially those of Hogan *et al.* (1986), with modification as described by Dony and Gruss (1987). Briefly, embryos were dissected free of extra-embryonic tissue at 10, 12 and 17 days p.c. and frozen directly on dry ice. 8 μm sections were cut at −20°C, dried for 10 min at 55°C on subbed slides (Gall and Pardue, 1971), and fixed in 4% formaldehyde. Sections were pretreated (Hafen *et al.*, 1983) in 2 × SSC for 30 min at 37°C, followed by pronase digestion, 0.1 mg/ml for 10 min. After rinsing in PBS, slides were fixed again in 4% paraformaldehyde, acetylated and dehydrated in graded ethanol.

RNA probes were labeled using SP6 or T7 polymerase and 100 μCi [^{35}S]-UTP (Amersham) essentially as described by the manufacturer (Promega Biotech). Labeled RNAs were precipitated with 10% trichloracetic acid and collected on nitrocellulose filters. Probes were eluted in 50 mM EDTA, 0.1% SDS at 65°C, and subjected to limited alkaline hydrolysis (Cox *et al.*, 1984). Probes were diluted to a specific activity of 5×10^4 c.p.m./μl in hybridization buffer containing: 50% formamide, 10% dextran sulfate, 0.3 M NaCl, 10 mM Tris, 10 mM $NaPO_4$, pH 6.8, 5 mM EDTA, 10 mM DTT, 1 × Denhardt's, and 10 μM S-ATP. Approximately 6−10 μl of hybridization mix were used, depending upon the size of the coverslip.

Hybridization was performed overnight at 50°C in a humid chamber. Slides were washed at 50°C in 50% formamide, 2 × SSC until coverslips floated free. Slides were digested with 20 mg/ml pancreatic RNase (Ingham *et al.*, 1985) and washed overnight at 37°C in 50% formamide, 2 × SSC. Following dehydration in graded ethanol, slides were dipped in Kodak NTB-2 emulsion, diluted 1:1 with water, and exposed for 6−14 days. Slides were developed and stained with Giemsa or cresyl violet acetate.

Acknowledgements

We thank Thomas Schowalter for excellent technical assistance, Bernd Föhring for computer analysis, and Rosemary Drescher for help in preparing and typing the manuscript. GD is the recipient of an Alexander von Humboldt Fellowship. Thanks are also due to B.Hogan for providing us with the cDNA library.

References

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408−1412.
Arriza,J., Weinberger,C., Cerreli,G., Glaser,T., Handelin,B., Housman,D. and Evans,R. (1987) *Science*, **237**, 268−275.
Baum,J.A., Geever,R. and Giles,N.H. (1987) *Mol. Cell. Biol.*, **7**, 1256−1266.
Boncinelli,E., Simeone,A., La Volpe,A., Faiella,A., Fidanza,V., Acampora,D. and Scotto,L. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 301−306.
Brown,R.S., Sander,C. and Argos,P. (1985) *FEBS Lett.*, **186**, 271−274.
Carroll,S.B. and Scott,M.P. (1985) *Cell*, **43**, 47−57.
Chirgwin,J.M., Przybyla,A.E., Macdonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294−5299.
Chowdhury,K., Deutsch,U. and Gruss,P. (1987) *Cell*, **48**, 771−778.
Colberg-Poley,A.M., Voss,S.D., Chowdhury,K. and Gruss,P. (1985) *Nature*, **314**, 713−718.
Cox,K., De Leon,D., Angrer,L. and Angerer,R. (1984) *Dev. Biol.*, **101**, 485−502.
Dony,C. and Gruss,P. (1987) *EMBO J.*, **6**, 2965−2975.
Fahrner,K., Hogan,B.L.M. and Flavell,R.A. (1987) *EMBO J.*, **6**, 1265−1271.
Frankel,A., Berg,J. and Pabo,C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4841−4845.
Friden,P. and Schimmel,P. (1987) *Mol. Cell. Biol.*, **7**, 2708−2717.
Gall,J. and Pardue,M. (1971) *Methods Enzymol.*, **38**, 470−480.
Garnier,J., Osguthorpe,D. and Robson,B. (1978) *J. Mol. Biol.*, **120**, 97−120.
Gaul,U., Seifert,E., Schuh,R. and Jäckle,H. (1987) *Cell*, **50**, 639−647.
Gehring,W.J. (1985) *Science*, **236**, 1245−1252.
Gehring,W. and Hiromi,Y. (1986) *Annu. Rev. Genet.*, **20**, 147−173.
Ginsberg,A., King,B. and Roeder,R. (1984) *Cell*, **39**, 479−489.
Green,S., Kumar,V., Krust,A., Walter,P. and Chambon,P. (1986) *Cold Spring Harbor Symp. Quant. Biol.*, **51**, 751−758.
Greene,G., Gilna,P., Waterfield,M., Baker,A., Hort,Y. and Shine,J. (1986) *Science*, **231**, 1150−1154.

Hafen,E., Levine,M., Garber,R. and Gehring,W. (1983) *EMBO J.*, **2**, 617–623.
Hanas,T.S., Hazuda,D.J., Bogenhagen,D.F., Wu,F. and Wu,C.W. (1983) *J. Biol. Chem.*, **258**, 14120–14125.
Hartshorne,T., Blumberg,H. and Young,E.T. (1986) *Nature*, **320**, 283–287.
Hiromi,Y., Kuroiwa,A. and Gehring,W.J. (1985) *Cell*, **43**, 603–613.
Hogan,B., Constantini,F. and Lacy,E. (1986) *Manipulating the Mouse Embryo. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Ingham,P., Howard,K. and Ish-Horowicz,D. (1985) *Nature*, **318**, 439–445.
Ingham,P., Ish-Horowicz,D. and Howard,K. (1986) *EMBO J.*, **5**, 1659–1665.
Jäckle,H., Tautz,D., Schuh,R., Seifert,E. and Lehman,R. (1986) *Nature*, **324**, 668–670.
Johnson,A.D. and Herskowitz,L. (1985) *Cell*, **42**, 237–247.
Johnston,M. (1987) *Nature*, **328**, 353–355.
Joyner,A., Kornberg,T., Coleman,K., Cox,D. and Martin,G. (1985) *Cell*, **43**, 29–37.
Kahn,P. and Graf,T. (eds) (1986) *Oncogenes and Growth Control*. Springer-Verlag, Heidelberg.
Kammerer,B., Guyonvarch,A. and Hubert,J.C. (1984) *J. Mol. Biol.*, **180**, 239–250.
Kozak,M. (1983) *Microbiol. Rev.*, **47**, 1–45.
Krumlauf,R., Holland,P., Mcvey,J. and Hogan,B. (1987) *Development*, **99**, 603–617.
Laughon,A. and Scott,M.P. (1984) *Nature*, **310**, 25–31.
Levine,M., Rubin,G.M. and Tjian,R. (1984) *Cell*, **38**, 667–673.
Lütcke,H.A., Chow,K.C., Mickel,F.S., Moss,K.A., Kern,H.F. and Scheele,G.A. (1987) *EMBO J.*, **6**, 43–48.
Maniatis,T., Fritsch,E. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Maxam,A. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
McGinnis,W., Hart,C., Gehring,W. and Ruddle,F. (1984) *Cell*, **38**, 675–680.
Messenguy,F., Dubois,E. and Descamps,F. (1986) *Eur. J. Biochem.*, **157**, 77–81.
Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609–1614.
Nüsslein-Vollhard,C. and Wieschhaus,E. (1980) *Nature*, **287**, 795–801.
Ollo,R. and Maniatis,T. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5700–5704.
Proudfoot,N. and Brownlee,G.G. (1976) *Nature*, **263**, 211–214.
Rijsewick,F., Schürmann,M., Wagenaar,E., Parren,P., Weigel,D. and Nusse,R. (1987) *Cell*, **50**, 649–657.
Rosenberg,U.B., Schröder,C., Preiss,A., Kienlin,A., Côté,S., Riede,E. and Jäckle,H. (1986) *Nature*, **319**, 336–339.
Ruiz i Altaba,A., Perry-O'Keefe,H. and Melton,D. (1987) *EMBO J.*, **6**, 3065–3070.
Salmeron,J., Jr and Johnston,S. (1986) *Nucleic Acids Res.*, **14**, 7767–7781.
Sap,J., Munoz,A., Damm,K., Goldberg,Y., Ghysdael,J., Leutz,A., Beug,H. and Vennström,B. (1987) *Nature*, **324**, 635–640.
Schuh,R., Aicher,W., Gaul,U., Côté,S., Preiss,A., Maier,D., Seifert,E., Nauber,U., Schröder,C., Kemler,R. and Jäckle,H. (1986) *Cell*, **47**, 1025–1032.
Tautz,D., Lehman,R., Schnürch,H., Schuh,R., Seifert,E., Kienlin,A., Jones,K. and Jäckle,H. (1987) *Nature*, **327**, 383–389.
Vincent,A., Colot,H. and Rosbash,M. (1985) *J. Mol. Biol.*, **186**, 149–166.
Weinberger,C., Thompson,C., Ong,E., Lebo,R., Gruol,D. and Evans,R. (1986) *Nature*, **324**, 641–646.

Received on February 15, 1988

Note added in proof

These sequence data will appear in the EMBL/GenBank/DDBJ nucleotide databases under the accession number Y00850.