Volume 304, number 2,3, 170–178 FEBS 11093 June 1992

Heterogeneous expression of four MAP kinase isoforms in human tissues

Fernando A. Gonzalez[a], David L. Raden[a,b], Mark R. Rigby[a,b] and Roger J. Davis[a,b]

[a]*Program in Molecular Medicine, Department of Biochemistry and Molecular Biology and* [b]*Howard Hughes Medical Institute, University of Massachusetts Medical School, Worcester, MA 01605, USA*

Received 10 February 1992; revised version received 10 April 1992

Mitogen-activated protein kinases (MAP kinases) are a group of closely related enzymes implicated in signal transduction pathways. We report the molecular cloning of four human proteins (p40^{mapk}, p41^{mapk}, p44^{mapk} and p63^{mapk}) with high homology to members of the MAP kinase family. Sequence analysis demonstrated that p44^{mapk} and p63^{mapk} were the products of distinct genes. However, the p40^{mapk} and p41^{mapk} were found to be related, and are likely to result from alternative processing of transcripts from a single gene. The heterogeneous expression of these human MAP kinase isoforms in different tissues may reflect the diversity of signal transduction pathways in differentiated cells.

Signal transduction; Molecular cloning

1. INTRODUCTION

Mitogen-activated protein kinases (MAP kinases) form a family of protein kinases that are activated by numerous extracellular stimuli in many cell types [1–7]. Members of this family of kinases have been implicated in a wide variety of cellular processes. MAP kinases are activated during the M phase of the meiotic cycle [6,8,9] and during mitogenic stimulation of quiescent cells [7,8,10]. It has been shown that MAP kinases can induce reorganization of microtubules in vitro [11], and also may modulate Myc- [12] and Jun- [13] mediated activation of gene expression. Thus, MAP kinases appear to have a fundamental role in multiple cellular processes that may share common signal transduction events [14].

It is believed that MAP kinases play a critical role in a protein kinase cascade pathway of signal transduction [10,15–18]. It has been suggested that MAP kinases represent 'switch' kinases that can transduce a tyrosine protein kinase activity into a serine/threonine protein kinase pathway [19–21]. This is because some MAP kinase isoforms are activated by phosphorylation on tyrosine and threonine [3,9,19–25]. However, the molecular mechanism of MAP kinase activation is unclear. It has been proposed that autophosphorylation plays a role in MAP kinase activation [23–25]. A MAP kinase activator has been identified and partially purified [16,23], and a kinase that phosphorylates MAP kinases may play a role in activation [17]. It is therefore possible that different cell types use related activation pathways in vivo to regulate and control MAP kinase activity [26].

Correspondence address: F.A. Gonzalez, Program in Molecular Medicine, University of Massachusetts Medical School, 373 Plantation Street, Worcester, MA 01605, USA.

The primary structure of peptide substrates recognized by members of this family of kinases has been examined [27,28] and the consensus for peptide substrate phosphorylation has been defined as Pro-Xaa$_n$-Ser/Thr-Pro (where Xaa is a neutral or basic amino acid and n = 1 or 2). Several proteins have been identified as substrates of MAP kinases. These include: microtubule-associated protein 2 [1,2], myelin basic protein [29], an S6 kinase [10,15,30], the epidermal growth factor receptor [7,31,32], the Myc and Jun proteins [12,13,33], and the kinase itself [21,23–25]. It has been proposed that MAP kinases have 'dual' specificity because they phosphorylate both serine/threonine and tyrosine residues [23–25]. However, it should be noted that the protein tyrosine kinase activity of MAP kinases has been documented only in autophosphorylation experiments, and evidence for the tyrosine phosphorylation of exogenous substrates has not been obtained [23–25].

Several MAP kinase cDNAs have been isolated from rat [4,22], *Xenopus* [8,34] and mouse [35]. Here we report the molecular cloning of four MAP-related protein kinases that are differentially expressed in human tissues. The heterogeneous expression of these MAP kinase isoforms suggests that signal transduction pathways in different tissues may be mediated by specific isoforms of these kinases.

2. EXPERIMENTAL

2.1. *Isolation and sequence analysis of p40^{mapk}, p41^{mapk}, p44^{mapk} and p63^{mapk} cDNAs*

The oligonucleotide pairs 5′-CAGTACATCGGCGAGGGC-3′ (sense)/5′-ATTGGAGGGCTTCAGGTC-3′ (antisense) and 5′-TACACCAAATCCATTGAC-3′ (sense)/5′-TGTCTCTTGGAAGATCAG-3′ (antisense) were used to amplify the sequences encompassing nucleotides 100–478 and 628–1,069, respectively, from the ERK1 cDNA using the polymerase chain reaction (PCR) and a rat brain

 Published by Elsevier Science Publishers B.V.

cDNA library (Stratagene Inc.) as template. PCR was performed using the GeneAmp kit (Perkin Elmer Cetus) with 35 cycles of amplification: denaturation at 94°C, 45 s; primer annealing at 55°C, 90 s; primer extension at 72°C, 20 s. The amplified products were subcloned into the *Sma*I site of pUC18 and sequenced to confirm their identity [4].

Plaques (2,000,000) from a HeLa cDNA library constructed in λZAP II vector were screened with the ERK1 probes according to the manufacturer's recommendations (Stratagene Inc.). After hybridization the library filters were washed under low stringency conditions (2 × SSC (15 mM sodium citrate, pH 7.0, 0.15 M NaCl), 0.1% SDS, 1 mM EDTA, 60°C). The phage were plaque-purified and the cDNAs were subcloned into the pBluescript II plasmid (Stratagene Inc.). DNA sequence analysis was performed using the dideoxynucleotide chain termination method using the Sequenase version 2.0 kit (US Biochemical Corp.) and oligonucleotide primers (Oligos Etc., Wilsonville, OR). Nested deletions to facilitate sequence analysis were made using exonuclease III and mung bean nuclease (Stratagene Inc.).

2.2. *Northern blot analysis*

The probes used for the Northern blot analysis were as follows: two probes were prepared from sequences present in both p41^{mapk} and p40^{mapk}: (i) a 672 bp *Eco*RI fragment corresponding to nucleotides 787–1458 of p41^{mapk} and 701–1,372 of p40^{mapk}; and (ii) a 558 bp *Sac*II–*Eco*RI fragment corresponding to nucleotides 229–786 of p41^{mapk} and 143–700 of p40^{mapk}. A specific probe for p41^{mapk} was prepared as a 229 bp *Eco*RI–*Sac*II fragment corresponding to nucleotides 1–228. A specific probe for p40^{mapk} was prepared as a 143 bp *Eco*RI–*Sac*II fragment corresponding to nucleotides 1–143. For p44^{mapk}, the oligonucleotides 5′-GAGGTGGAGATGGTG-3′ (sense) and 5′-TCCCGGAGCGTGCGC-3′ (antisense) were used for PCR amplification of nucleotides 7–191 of the p44^{mapk} cDNA. For p63^{mapk}, the oligonucleotides 5′-AGCGAGGTACAGCGC-3′ (sense) and 5′-GGCTGGCGTGCTCTG-3′ (antisense) were used for PCR amplification of nucleotides 1,673–2058 of the p63^{mapk} cDNA.

Northern blots were performed using 2 μg of poly(A)$^+$ mRNA isolated from different human tissues and cultured HeLa cells, fractionated by denaturing agarose gel electrophoresis and transferred onto a nylon membrane (Clontech, #7760-1 and #6522-2). The blots were hybridized with each probe in 5 × SSPE (10 mM NaH_2PO_4, pH 7.4, 0.15 M NaCl, 1 mM EDTA), 10 × Denhardt's solution, 100 μg/ml salmon sperm DNA, 50% formamide, 2% sodium dodecyl sulfate at 42°C. After hybridization the blots were washed in 2 × SSC, 0.05% sodium dodecyl sulfate, 1 mM EDTA at room temperature and autoradiographed.

2.3. *Radioactive labeling of DNA probes*

DNA probes were labeled by random priming using the Multiprime DNA Labelling System and [α-^{32}P]dCTP (Amersham).

2.4. *Nucleotide sequence accession number*

The sequences of p41^{mapk}, p40^{mapk}, p44^{mapk} and p63^{mapk} reported here will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers Z11694, Z11695, Z11696 and X59727, respectively.

3. RESULTS

3.1. *Molecular cloning of four human enzymes related to MAP kinases*

Two probes based on the rat *ERK1* cDNA [4] were used to screen a human HeLa cDNA library (cloned in λZAP II vector) at low stringency (see section 2). Multiple hybridizing cDNA clones were initially identified (18 phages). Upon secondary screening and further characterization of the clones it was found that 13 contained cDNAs encoding four different proteins that are homologous to other members of the MAP kinase fam-

A: p41^{mapk}

```
                                                                                                                AATTCGAGGAATGGTCCAGTGGTGT    25
CGGCTCTTCAGCTCTCGCTCGCGTCTTCGTTCCTCTCCCGGTCAGCGTCGGCGGCTGCACCGGCGGCGGGGCAGTCTGCGGCAGGGGCGACAAGAGCTGAGGCGCGGCCGCCGAGCGTCGAGCTCAGCGCGGCGGAGGCGGCGGCCCGGCAGCCAAC  104

ATG GCG GCG GCG GCG GCG GCG GGC GCG GGC CCG GAG ATG GTC CGC GGG CAG GTG TTC GAC GTG GGG CCG CGC TAC ACC AAC CTC TCG TAC ATC GGC GAG GGC GCC TAC GGC ATG GTG TGC  304
Met Ala Ala Ala Ala Ala Gly Ala Gly Pro Glu Met Val Arg Gly Gln Val Phe Asp Val Gly Pro Arg Tyr Thr Asn Leu Ser Tyr Ile Gly Glu Gly Ala Tyr Gly Met Val Cys  (40)

TCT GCT TAT GAT AAT GTC AAC AAA GTT CGA GTA GCT ATC AAG AAA ATC AGC CCC TTT GAG CAC CAG ACC TAC TGC CAG AGA ACC CTG AGG GAG ATA AAA ATC TTA CTG CGC TTC AGA CAT  424
Ser Ala Tyr Asp Asn Val Asn Lys Val Arg Val Ala Ile Lys Lys Ile Ser Pro Phe Glu His Gln Thr Tyr Cys Gln Arg Thr Leu Arg Glu Ile Lys Ile Leu Leu Arg Phe Arg His  (80)

GAG AAC ATC ATT GGA ATC AAT GAC ATT ATT CAA GCA CCA ACC ATC GAG CAA ATG AAA GAT GTA TAT ATA GTA CAG GAC CTC ATG GAA ACA GAT CTT TAC AAG CTC TTG AAG ACA CAA CAC  544
Glu Asn Ile Ile Gly Ile Asn Asp Ile Ile Gln Ala Pro Thr Ile Glu Gln Met Lys Asp Val Tyr Ile Val Gln Asp Leu Met Glu Thr Asp Leu Tyr Lys Leu Leu Lys Thr Gln His  (120)

CTC AGC AAT GAC CAT ATC TGC TAT TTT CTC TAC CAG ATC CTC AGA GGG TTA AAA TAT ATC CAT TCA GCT AAC GTT CTG CAC CGT GAC CTC AAG CCT TCC AAC CTG CTG CTC AAC ACC ACC  664
Leu Ser Asn Asp His Ile Cys Tyr Phe Leu Tyr Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu His Arg Asp Leu Lys Pro Ser Asn Leu Leu Leu Asn Thr Thr  (160)

TGT GAT CTC AAG ATC TGT GAC TTT GGC CTG GCC CGT GTT GCA GAT CCA GAC CAT GAT CAC ACA GGG TTC CTG ACA GAA TAT GTG GCC ACA CGT TGG TAC AGG GCT CCA GAA ATT ATG TTG  784
Cys Asp Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Val Ala Asp Pro Asp His Asp His Thr Gly Phe Leu Thr Glu Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Ile Met Leu  (200)

AAT TCC AAG GGC TAC ACC AAG TCC ATT GAT ATT TGG TCT GTA GGC TGC ATT CTG GCA GAA ATG CTT TCT AAC AGG CCC ATC TTT CCA GGG AAG CAT TAT CTT GAC CAG CTG AAC CAC ATT  904
Asn Ser Lys Gly Tyr Thr Lys Ser Ile Asp Ile Trp Ser Val Gly Cys Ile Leu Ala Glu Met Leu Ser Asn Arg Pro Ile Phe Pro Gly Lys His Tyr Leu Asp Gln Leu Asn His Ile  (240)

TTG GGT ATT CTT GGA TCC CCA TCA CAA GAA GAC CTG AAT TGT ATA ATA AAT TTA AAA GCT AGG AAC TAT TTG CTT TCT CTT CCA CAC AAA AAT AAG GTG CCA TGG AAC AGG CTG TTC CCA 1024
Leu Gly Ile Leu Gly Ser Pro Ser Gln Glu Asp Leu Asn Cys Ile Ile Asn Leu Lys Ala Arg Asn Tyr Leu Leu Ser Leu Pro His Lys Asn Lys Val Pro Trp Asn Arg Leu Phe Pro  (280)

AAT GCT GAC TCC AAA GCT CTG GAC TTA TTG GAC AAA ATG TTG ACA TTC AAC CCA CAC AAG AGG ATT GAA GTA GAA CAG GCT CTG GCC CAC CCA TAT CTG GAG CAG TAT TAC GAC CCG AGT 1144
Asn Ala Asp Ser Lys Ala Leu Asp Leu Leu Asp Lys Met Leu Thr Phe Asn Pro His Lys Arg Ile Glu Val Glu Gln Ala Leu Ala His Pro Tyr Leu Glu Gln Tyr Tyr Asp Pro Ser  (320)

GAC GAG CCC ATC GCC GAA GCA CCA TTC AAG TTC GAC ATG GAA TTG GAT GAC TTG CCT AAG GAA AAG CTC AAA GAA CTA ATT TTT GAA GAG ACT GCT AGA TTC CAG CCA GGA TAC AGA TCT 1264
Asp Glu Pro Ile Ala Glu Ala Pro Phe Lys Phe Asp Met Glu Leu Asp Asp Leu Pro Lys Glu Lys Leu Lys Glu Leu Ile Phe Glu Glu Thr Ala Arg Phe Gln Pro Gly Tyr Arg Ser  (360)

TAA ATTTGTCAGGACAAGGGCTCAGAGGACTGGACGTGCTCAGACATCGGTGTTCTTCTTCCCAGTTCTTGACCCCTGGTCCTGTCTCCAGCCCGTCTTGGCTTATCCGCTTTGACTCCTTTGAGCCGTTTGGAGGGGCGGTTTCTGGTAGTTGTGCC 1422
***

TTTTATGCTTTCAAAGAATTTCTTCAGTCCAGAGAATTCCTCCTGGCAGCCCTGTGTGTGTCACCCATTGGTGACCTGCGGGCAGTATGTACTTCAGTGCACCTTACTGCTTACTGTTGCTTTAGTCACTAATTGCTTTCTGCTTTGAAAGATGCAGTGG 1581
TTCCTCCCTCTCGTGAATCGTTTTCTACATGATGCCCTGCTGACCATGCAGCCGCACCAGAGACAGATTCTTCCCCAATTGGCTCTAGTCACTGGCATCTCACTTTATGATAGGGAAGGCTACTACCTAGGGCACTTTAAGTCAGTGACAGCCCCTTAT 1740
TTGCACTTCACCTTTTGACCATAACTGTTTCCCCAGAGCAGGAGCTTGTGGAAATACCTTGGCTGATGTTGCAGCCTGCAGCAAGTGCTTCCGTCTCCGGAATCCTTGGGGAGAGCACTTGTCCACGTCTTTTCTCATATCATGGTAGTCACTAACATATA 1899
TAAGGTATGTGCTATTGGCCCAGCTTTTAGAAAATGCAGTCATTTTTCTAAATAAAAAAGGAAGTACTGCACCCAGCAGTGTCACTCTGTAGTTACTGTGGTCACTTGTACCATATAGAGGTGTAACACTTGTCAAGAAGCGTTATGTGCAGTACTTAA 2058
TGTTTGTAAGACTTACAAAAAAAAGATTTAAAGTGGCAGCTTCACTCGACATTTGGTGAGAGAAGTACAAAGGTTGCAGTGCTGAGCTGTGGGCCGTTTCTGGGGCATGTCCCAGGGTGAACTCCACATGCTGGTGCATATACGCCCTTGAGCTACTTCA 2217
AATGTGGTTTATACCTCGCAGATACAAGAATCTTTATGAATATACAATTCTTTTTCCTTCTACAGCTTAGCTCCGTCTTTTCAACCACGAACATTTAAAACCCGACCTACTAGCACTGTTCTGTCCTCAAGTAC  2351
```

Fig. 1A. Nucleotide sequence of p41^{mapk} cDNA and its deduced protein sequence. An in-frame termination codon in the sequence 5′ to the initiation codon is underlined. The termination codon for the kinase transcript is represented by three asterisks (***). Protein sequence sub-domains conserved in protein kinases [37] are double-underlined. The major open reading frame present in this cDNA encodes a predicted protein kinase with a calculated molecular weight of 41 kDa. Gln^{91} is circled (see section 3 for details). The autophosphorylation site proposed to be required for kinase activation (TEY) is boxed [21,23–25].

B: p40^{mapk}

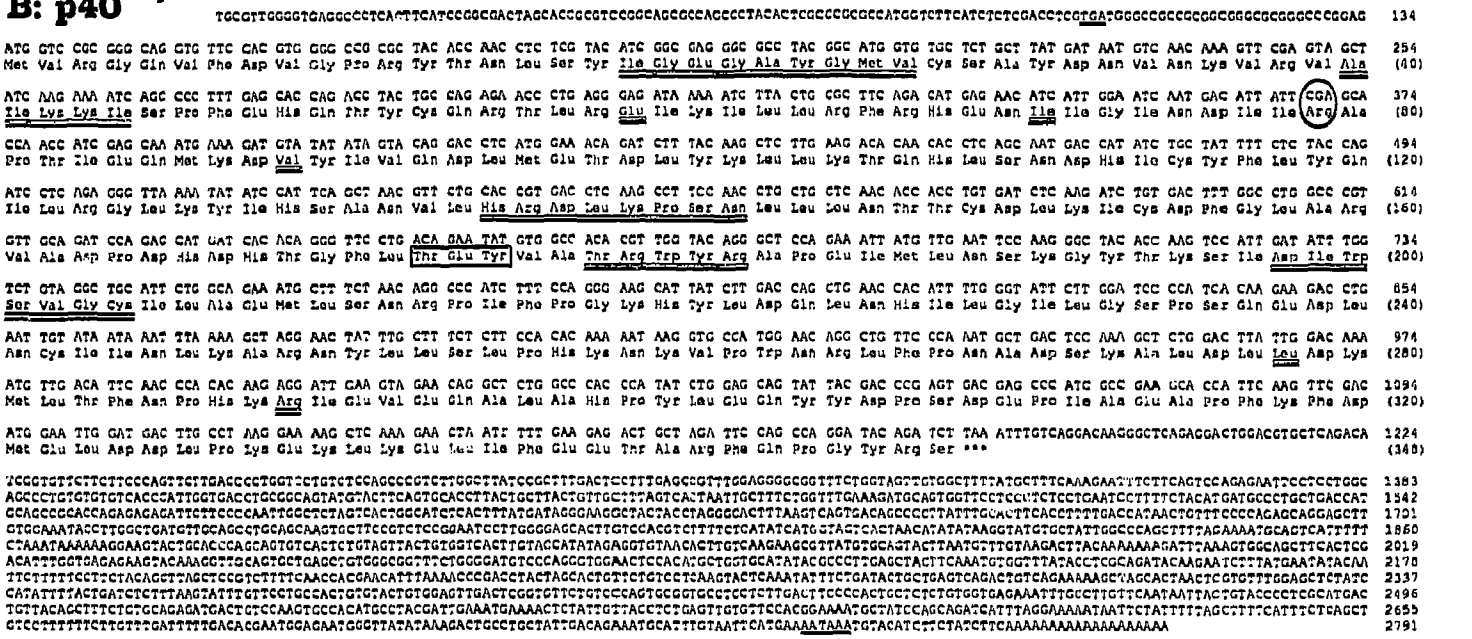

Fig. 1B. Nucleotide sequence of p40^{mapk} cDNA and its deduced protein sequence. An in-frame termination codon in the sequence 5′ to the initiation codon is underlined. The termination codon for the kinase transcript is represented by three asterisks (***). Protein sequence sub-domains conserved in protein kinases [37] are double-underlined. The major open reading frame present in this cDNA encodes a predicted protein kinase with a calculated molecular weight of 40 kDa. Arg^{79} is circled (see section 3 for details). The autophosphorylation site (TEY) is boxed. A poly-adenylation sequence (AATAAA) in the 3′ untranslated region is underlined.

ily. These kinases were designated p41^{mapk}, p40^{mapk}, p44^{mapk} and p63^{mapk} (Fig. 1).

3.2. *p40^{mapk} and p41^{mapk} are highly related protein kinases*

After examination of the sequence of cDNAs from 13 positive clones, we discovered that 7 had an open reading frame encoding a 41 kDa protein with homology to MAP kinases. This kinase was named p41^{mapk} (Fig. 1A). Comparison of the sequence of p41^{mapk} with other members of the MAP kinase family (Table I) indicated that a high degree of homology exists between p41^{mapk} and

C: p44^{mapk}

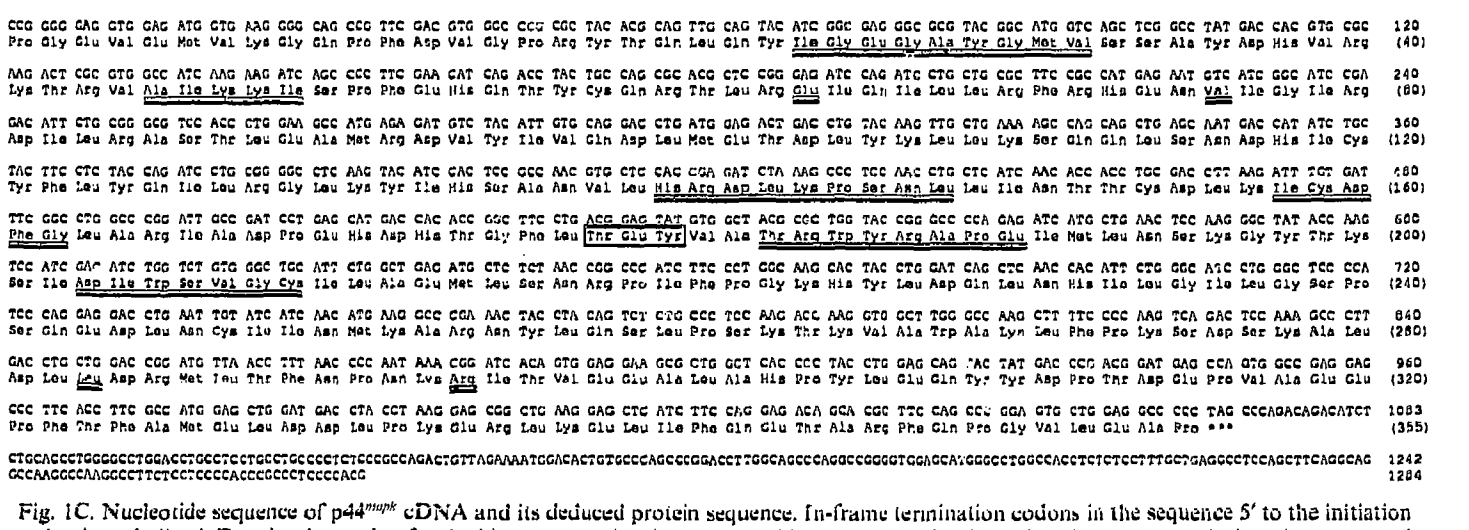

Fig. 1C. Nucleotide sequence of p44^{mapk} cDNA and its deduced protein sequence. In-frame termination codons in the sequence 5′ to the initiation codon is underlined. Termination codon for the kinase transcripts is represented by three asterisks (***). Protein sequence sub-domains conserved in protein kinases [37] are double-underlined. The major open reading frame present in this cDNA encodes a predicted protein kinase with an estimated molecular weight of 44 kDa. The autophosphorylation site (TEY) is boxed. Similar to the cDNA isolated for ERK1 [4] this cDNA is not full-length and is truncated at the 5′ end.

D: p63^{mapk}

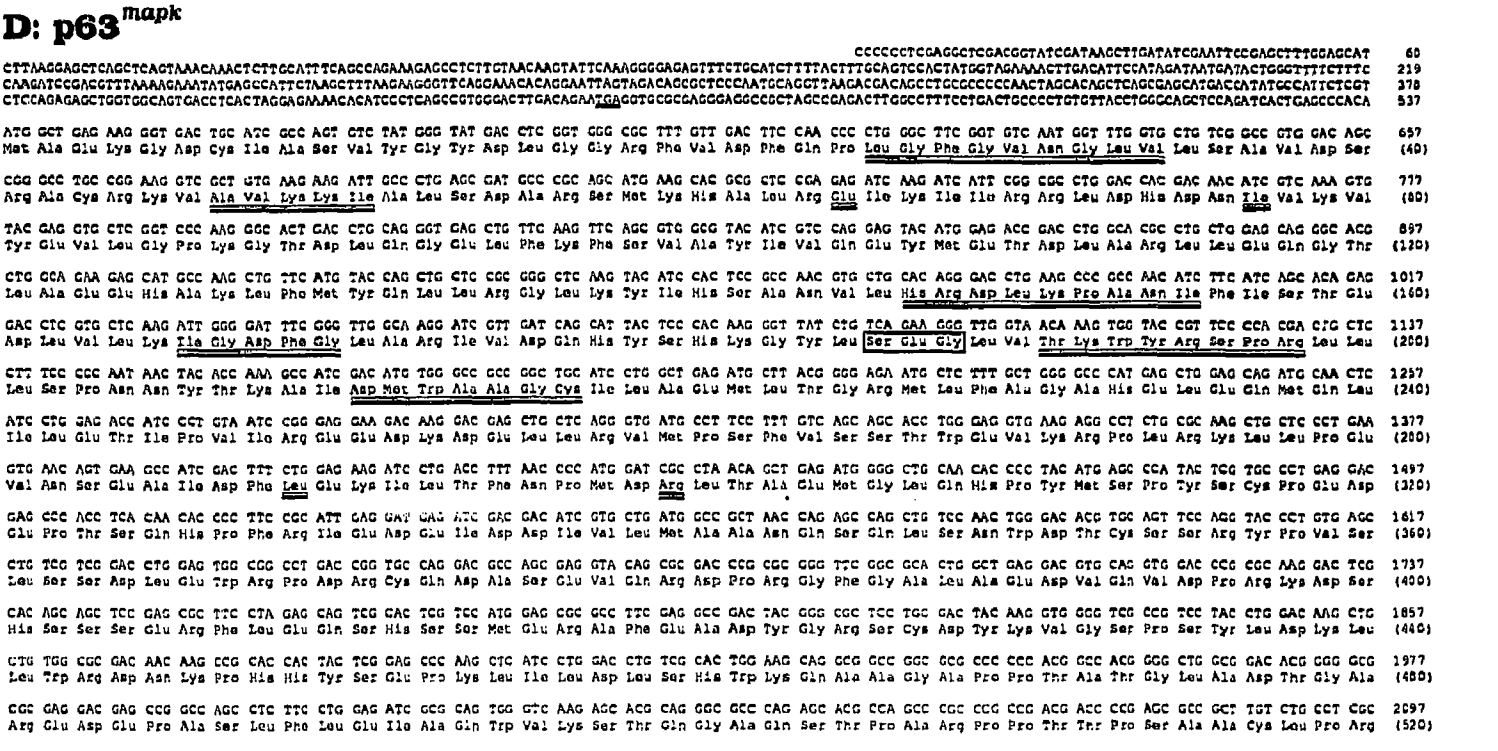

Fig. 1D. Nucleotide sequence of p63^{mapk} cDNA and its deduced protein sequence. An in-frame termination codon in the sequence 5′ to the initiation codon is underlined. The termination codon for the kinase transcripts is represented by three asterisks. Protein sequence sub-domains conserved in protein kinases [37] are double-underlined. The major open reading frame present in this cDNA encodes a predicted protein kinase with a calculated molecular weight of 63 kDa. The protein kinase p63^{mapk} contains the sequence SEG (boxed) in place of the TEY autophosphorylation motif present in other members of the MAP kinase family. An in-frame stop codon (TGA) is located 90 bp 5′ of the predicted ATG translational initiation codon. Within the long 5′ untranslated region there are an additional six potential translational initiation codons. These ATG codons are not located in a favorable context for initiation [36] and correspond to short open reading frames. The significance of the multiple ATG codons in the 5′ region of the sequence is discussed in the text.

the mouse mitogen activated kinase, pp42/MAP [35], the rat extracellular signal-regulated kinase 2, ERK2 [22] and the *Xenopus* M phase MAP kinase, Xp42/MPK1 [8,34].

A second group of clones (2 phages) contain a cDNA with an open reading frame encoding a protein with a predicted molecular weight of 40 kDa. This enzyme was named p40^{mapk} (Fig. 1B). The protein kinases p41^{mapk} and p40^{mapk} share a high degree of homology, namely 96.9% identical (Table I). In fact these two kinases differ only at their amino-termini and a single amino acid in the kinase domain. The divergent sequences in the 5′ region of these cDNAs are illustrated in Fig. 2. The amino acid residues 13–360 of p41^{mapk} are identical to amino acid residues 1–348 of p40^{mapk} with the exception that Gln^{91} in p41^{mapk} is replaced with Arg (residue 79) in p40^{mapk}. The cDNA encoding p40^{mapk} has a putative polyadenylation signal (AATAAA) and a poly-A tail at the 3′ end (Fig. 1B).

In order to distinguish the mRNAs encoding p41^{mapk} and p40^{mapk} from each other, we performed Northern blot analysis of poly(A)$^+$ mRNA from HeLa cells (Fig. 3). Probes from regions of identity between p41^{mapk} and p40^{mapk} (namely, nucleotides 229–786 and 787–1458 of p41^{mapk}) hybridized with three mRNAs of 5.5, 3.3 and 2.2 kb. In addition, a small (< 1.0 kb) transcript hybridized with one of the probes (nucleotides 787–1458 of p41^{mapk}). A DNA probe that is specific for p40^{mapk} (nu-

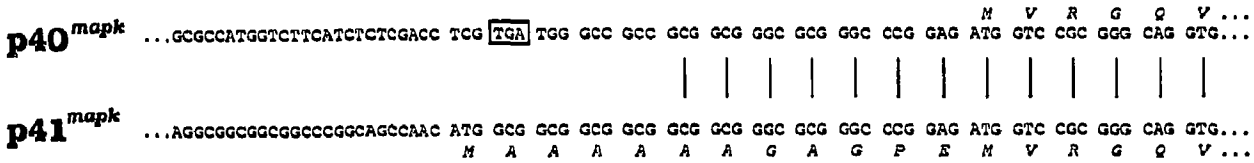

Fig. 2. Comparison of the 5′ region of p40^{mapk} and p41^{mapk}. The region of near identity of these two sequences corresponds to nucleotides 200–2,351 for p41^{mapk} and 113–2,264 for p40^{mapk}. The divergent 5′ region of these cDNAs is illustrated. Identity sequences are illustrated by vertical lines. An in-frame stop codon p40^{mapk} is boxed.

cleotides 1–228) hybridized with the 5.5 kb mRNA (Fig. 3). In contrast a probe specific for p42^{mapk} (nucleotides 1–143) hybridized with two mRNAs of 5.5 and 2.2 kb (Fig. 3). Therefore, it is likely that the mRNA for p41^{mapk} is a 5.5 kb transcript, while p40^{mapk} may be expressed from two transcripts of different sizes (2.2 and 5.5 kb).

The 5′-untranslated region of p40^{mapk} has a high degree of homology with human acidic ribosomal phosphoprotein P1 (~ 0.5 kb mRNA) [44]. Thus a small abundant transcript (≪ 1.0 kb) hybridized with the p40^{mapk} probe (nucleotides 1–143). This small transcript did not hybridize with any of the other MAP kinase probes. An actin probe hybridized to a 2.0 kb transcript (Fig. 3), indicating that the mRNA preparation used for this analysis was not degraded.

3.3. *p44*^{mapk} *protein kinase*

An additional set of clones (2 phages) contain a cDNA encoding a protein kinase. This kinase was named p44^{mapk} (Fig. 1C) because of its high degree of homology (98.3%) with a MAP-related protein kinase with molecular weight of 44 kDa, rat extracellular signal-regulated kinase 1, ERK1 [4] (see Table I). Similar to the isolated ERK1 cDNA [4] the cDNAs identified for p44^{mapk} were truncated at the 5′ end (Fig. 1C). The protein kinase p44^{mapk} is closely related to p41^{mapk} and p40^{mapk} (84.7 and 85.9% identity, respectively; see Table I).

3.4. *p63*^{mapk} *protein kinase*

A fourth class of cDNA was found in 2 phages. This type of cDNA contains an open reading frame encoding a 63 kDa protein (named p63^{mapk}, Fig. 1D) with homology (54.6% identity) to the rat extracellular signal-regulated kinase 3, ERK3 [22]. An in-frame stop codon was found 5′ of the predicted translational initiation codon; however, an additional six ATG codons are located within the long 5′ untranslated region. These potential initiation codons (corresponding to short open reading frames) do not conform to the consensus sequence for translational initiation codons [36].

A significant similarity between p63^{mapk} and ERK3 is the substitution of the Ala-Pro-Glu motif in the conserved kinase sub-domain VIII [37] with the sequence Ser-Pro-Arg. However, although p63^{mapk} is most closely related to ERK3 careful examination of an alignment of the sequences of these two kinases (Fig. 4) reveals that the highest homology is present within the predicted kinase domain (71.8% identity) while the COOH-terminal region is markedly different (only 26.8% identity). The overall identity between p63^{mapk} and ERK3 is 54.6%. The protein kinase p63^{mapk} is also homologous to the other rat MAP-related kinases ERK1 (41.4% identity) and ERK2 (41.6% identity), the mouse mitogen activated protein kinase pp42/MAP (41.1% identity) and the *Xenopus* M phase MAP kinase Xp42/MPK1 (41.0% identity). Other members of the MAP kinase family (KSS1 and FUS3 [38,39], CDK2 [40–42] and human CDC2 protein kinase [43]) are less closely related to p63^{mapk} (see Table I).

3.5. *Tissue distribution of MAP-related protein kinases*

Northern blot analysis with a probe that recognizes both p41^{mapk} and p40^{mapk} revealed that there are at least three mRNAs (2.8, 4.6 and 5.5 kb) present in human tissues that could encode for these kinases (Fig. 5A). The multiplicity of mRNAs hybridizing to the p40/41^{mapk} probe suggests that these represent very closely related transcripts. There is heterogeneity in the tissue distribution of the three mRNAs that hybridized to the p40/41^{mapk} probe. The largest transcript (5.5 kb) was expressed at highest levels in brain and at high levels in heart and lung, but was not detected in pancreas. The 4.6 kb transcript was most abundant in skeletal muscle, heart muscle and liver. The smallest transcript (2.8 kb) was most abundant in skeletal muscle, heart muscle and brain. This pattern of expression may be indicative of specialized tissue-specific signal transduction pathways that utilize different isoforms of this highly related subclass of MAP kinases.

In contrast to the multiple mRNAs that hybridized to the p40/41^{mapk} probe it was observed that only a single major transcript was detected by Northern analysis using probes for p44^{mapk} and p63^{mapk}. The p44^{mapk} probe identified a 2.2 kb mRNA that was expressed at high levels in lung tissue (Fig. 5B). The p63^{mapk} probe identified a 5.5 kb mRNA that was very abundant in heart and brain tissues (Fig. 5C).

Examination of the Northern blot hybridization patterns revealed that heart, brain and lung tissues expressed transcripts for the four human MAP kinases. However, in contrast to heart muscle, skeletal muscle only expressed significant levels of transcripts for p40^{mapk}/41^{mapk} and p44^{mapk}. Similarly, placenta, liver and kidney tissues expressed transcripts for p40^{mapk}/p41^{mapk}

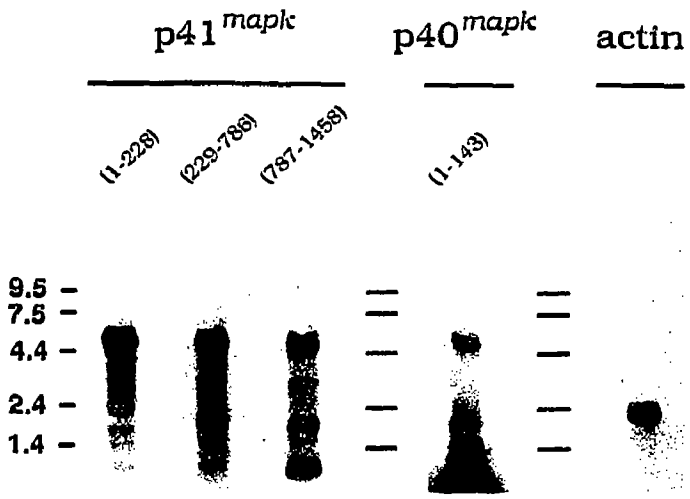

Fig. 3. Identification of mRNAs encoding $p41^{mapk}$ and $p40^{mapk}$ in HeLa cells. Northern blots containing 2 μg of poly(A)$^+$ mRNA isolated from HeLa cells were hybridized with probes for $p41^{mapk}$, $p40^{mapk}$ and actin. The probes used are indicated in parentheses as nucleotide numbers. Nucleotides 229–786 of $p41^{mapk}$ correspond to 143–700 of $p40^{mapk}$, nucleotides 787–1,458 of $p41^{mapk}$ correspond to nucleotides 701–1,372 of $p40^{mapk}$. RNA size markers (in kb) are shown to the left.

and $p44^{mapk}$. On the other hand, pancreas expressed only low levels of a transcript that hybridized to the $p40/p41^{mapk}$ probe and did not express detectable levels of $p44^{mapk}$ or $p63^{mapk}$ mRNAs.

Northern blot hybridization with an actin probe revealed that the mRNA preparations were not degraded (Fig. 5D). A ~2.0 kb mRNA was detected in all tissues. An additional actin mRNA (~1.8 kb) was detected in heart and skeletal muscle samples [45]. The pancreas showed the lowest levels of MAP-kinase mRNAs because a substantial fraction of the total mRNA isolated from this tissue encodes two small transcripts, proinsulin and α-amylase.

4. DISCUSSION

This study describes the sequences of four human kinases with homology to the MAP family of protein kinases. The existence of a larger family of MAP kinases is supported by the multiplicity of transcripts that specifically hybridized to human MAP kinase probes during Northern analysis (Fig. 5). It is therefore likely that there may be additional MAP kinase-related enzymes in human tissues than the four described here. Indeed the existence of an extended family of MAP kinase isoforms has previously been proposed by Boulton et al. [22].

Table I

Comparison between $p41^{mapk}$, $p40^{mapk}$, $p44^{mapk}$, $p63^{mapk}$ and related kinases

	Percent identity			
	$p41^{mapk}$ (1–360)	$p40^{mapk}$ (1–348)	$p44^{mapk}$ (1–355)	$p63^{mapk}$ (1–341)
$p41^{mapk}$	100.0			
$p42^{mapk}$	96.9	100.0		
$p44^{mapk}$	84.7	85.9	100.0	
$p63^{mapk}$	39.2	37.9	41.6	100.0
pp42/MAP	98.6	96.8	86.0	41.1
ERK1	84.4	82.2	98.3	41.4
ERK2	98.3	96.8	86.0	41.6
ERK3	41.4	40.8	44.9	72.7
Xp42/MPK1	95.0	92.2	85.1	41.0
KSS1	49.4	51.4	48.9	34.9
FUS3	48.9	51.1	50.0	30.8
CDK2	31.6	35.6	36.2	29.6
hCDC2	28.9	33.3	31.2	26.4

Percentage of amino acid identity between the protein kinase sequences was calculated from computer-generated alignments (MacVector Computer Analysis Software, International Biotechnologies Inc., New Haven, CT). Alignments were made over the residues indicated in parentheses. The protein kinases used in these alignments are: mouse mitogen activated protein kinase, pp42/MAP [35]; rat extracellular signal-regulated kinase 1, ERK1 [4]; rat extracellular signal-regulated kinase 2, ERK2 [22]; rat extracellular signal-regulated kinase 3, ERK3 [22]; *Xenopus* M phase MAP kinase, Xp42/MPK1 [8,34]; *S. cerevisiae* serine/threonine kinases mediating yeast responses to pheromones, KSS1 and FUS3 [38,39]; human cell division kinase 2, CDK2 [40–42]; and human CDC2 protein kinase, hCDC2 [43].

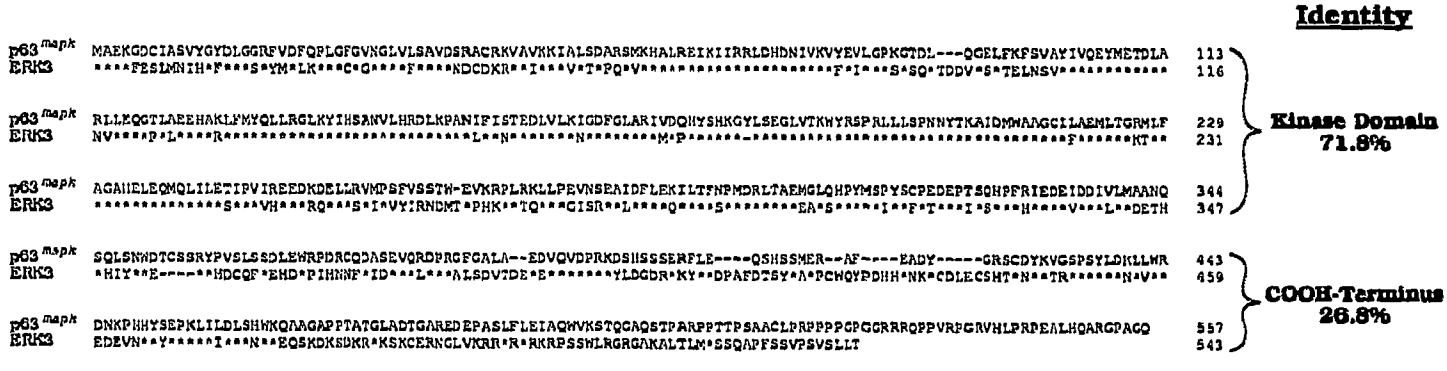

Fig. 4. Comparison of the protein sequence of $p63^{mapk}$ with ERK3. Sequence alignment was computer-generated (MacVector Computer Analysis Software, International Biotechnologies Inc., New Haven, CT) and visually optimized. Asterisks indicate identity to the ERK3 sequence; dashes indicate spaces introduced to optimize sequence alignment. Percent identities for the kinase domain (residues 1–341 in $p63^{mapk}$) and COOH-terminal domain (residues 342–557 in $p63^{mapk}$) are presented. The overall sequence identity between $p63^{mapk}$ and ERK3 is 54.6%.

4.1. *$p40^{mapk}$ and $p41^{mapk}$ differ only at their NH_2-termini and a single amino acid in the kinase domain*

Careful examination of the sequences of the $p41^{mapk}$ cDNA and the $p40^{mapk}$ cDNA reveals that the nucleotide sequences are very similar (Fig. 1). In fact nucleotides 200–2351 in the $p41^{mapk}$ cDNA are identical to nucleotides 113–2264 of the $p40^{mapk}$ cDNA except for a single nucleotide replacement. The region of identity between the two sequences includes most of the open reading frame and all of the 3′ un-translated region of the $p41^{mapk}$ cDNA (this cDNA does not contain a poly-A tail). In contrast, the 5′ region of these cDNAs are different. This results in distinct initiation codons and NH_2-termini of the two protein kinases (Fig. 2). It is possible that the divergent 5′ sequences of these two cDNAs are the result of alternative splicing of two different exons corresponding to the 5′ untranslated region and the NH_2-terminus to a common set of 3′ exons. Previously, this form of alternative expression of isoforms has been described for the c-*abl* protein tyrosine kinase [46]. The c-*abl* gene has alternative promoters with 5′ exons that are spliced onto common 3′ exons to give rise to two different enzymes, type I and IV kinases that differ only at their NH_2-terminal region [46]. Another example is represented by the human p70 S6 kinase [47]. The cDNA of two forms of the p70 S6 kinase have been isolated that encode divergent 5′ regions with alternative potential initiation codons while most of the open reading frame of the two cDNAs is identical [47].

In addition to the divergent NH_2-termini of $p41^{mapk}$ and $p40^{mapk}$, there is an additional single nucleotide change that replaces Glu^{91} (in $p41^{mapk}$) with Arg (residue 79 in $p40^{mapk}$). This difference could either be: (i) evidence for additional alternative splicing or editing of a single RNA transcript; (ii) evidence for the presence of a family of related genes that separately encode $p41^{mapk}$ and $p40^{mapk}$; (iii) caused by a sequence difference in two expressed alleles of the same gene in HeLa cells; or (iv) a cloning artifact. The elucidation of the true genetic relationship between $p41^{mapk}$ and $p40^{mapk}$ must await genomic cloning of these kinases.

4.2. *$p63^{mapk}$ is a novel member of the MAP kinase family*

$p41^{mapk}$ and $p40^{mapk}$ have a high degree of homology with rat ERK2 [22], mouse pp42/MAP [35] and *Xenopus* Xp42/MPK1 [8,34] (see Table I). It is likely that these enzymes are phylogenetic homologs from different species. The protein kinase $p44^{mapk}$ is highly related to ERK1 and these enzymes may therefore also represent phylogenetic homologs. However, although the protein kinase $p63^{mapk}$ is most closely related to ERK3, there is only a 54.6% overall identity between these two enzymes (Fig. 4). Because of the relatively low level of identity it is possible that $p63^{mapk}$ and ERK3 are related enzymes that belong to a sub-family, but may not be true phylogenetic homologs.

The enzymes $p63^{mapk}$ and ERK3 share a unique feature in that they contain the sequence Ser-Pro-Arg in substitution for the Ala-Pro-Glu motif present in the conserved protein kinase sub-domain VIII [37]. In a recent comparison of 154 eukaryotic protein kinases [48] all but one contained a Glu residue within the APE motif (the *S. cerevisiae* PHO85 protein kinase has a conservative substitution of the Glu with Asp). The replacement of Glu with Arg in $p63^{mapk}$ and ERK3 is therefore a distinctive non-conservative change within the protein kinase sub-domain VIII. In contrast, the replacement of Ala in the APE motif with Ser has been found in several other protein kinases [48].

$p63^{mapk}$ and ERK3 differ from most MAP kinases in that the autophosphorylation motif Thr-Glu-Tyr, proposed to be involved in kinase activation, is substituted by Ser-Glu-Gly (see Figs. 1D and 4). This observation raises important mechanistic questions about the physiological regulation of this sub-family of MAP kinases. Investigations of the possible involvement of the SEG sequence in the activation of $p63^{mapk}$ and the examination of alternative regulatory mechanisms of kinase activation are warranted.

The presence of multiple potential initiation codons in the 5′ region of the $p63^{mapk}$ open reading frame is intriguing. These upstream ATG codons may reflect an intron-containing mRNA [49]. On the other hand, this 'ATG-burdened' 5′ region may impair translation of the $p63^{mapk}$ mRNA in ectopic tissues thereby providing additional regulatory control on the expression of this gene [49].

4.3. *Tissue distribution of human MAP kinases*

Different isoforms of MAP kinases may play diverse roles in signal transduction pathways in different tissues. Our experiments show that the distribution and relative abundance of MAP kinase transcripts are not uniform in human tissues (Fig. 5). Thus, the $p63^{mapk}$ mRNA is more abundant in heart muscle and brain than in any other human tissue tested. On the other hand, the $p44^{mapk}$ mRNA is expressed at high levels in lung tissue, which is rich in fibroblasts, epithelial and endothelial cells.

It is interesting that the three major transcripts (2.8, 4.6 and 5.5 kb) that hybridize to the $p40^{mapk}$/$p41^{mapk}$ probe do not display the same pattern of expression in all human tissues investigated. Brain, placenta, lung and kidney tissues express a high level of the largest transcript. Heart and skeletal muscle contain all three transcripts, while liver is richer in the 4.6 kb transcript. The multiplicity of transcripts hybridizing to the $p40^{mapk}$/$p41^{mapk}$ probe may represent: (i) a subfamily of very closely related genes; (ii) alternatively spliced forms of a single transcript, or (iii) expression of transcripts created from different promoters spliced onto common 3′ exons.

The discrete pattern of expression of human MAP

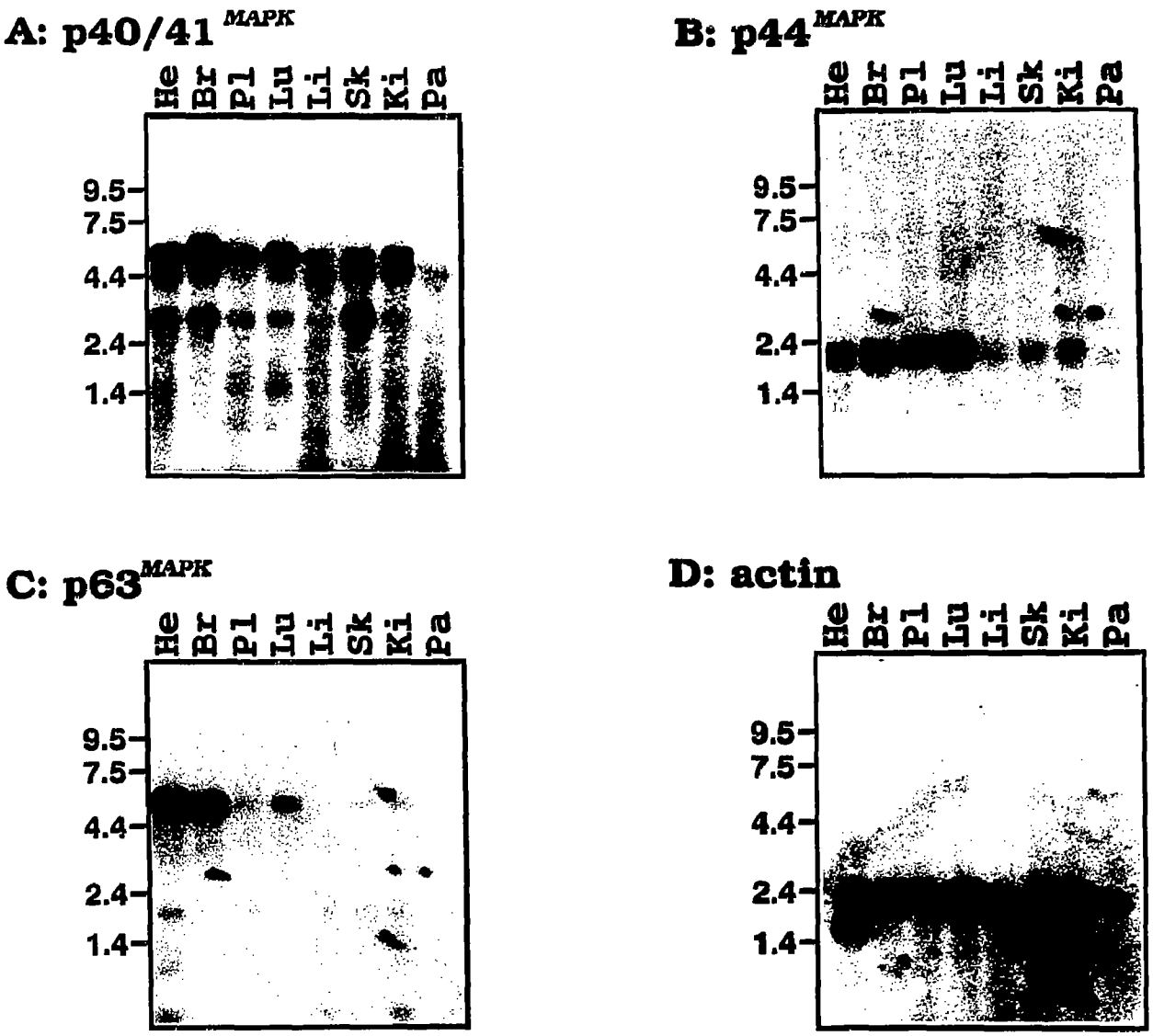

Fig. 5. Distinct patterns of expression of p40^{mapk}, p41^{mapk}, p44^{mapk} and p63^{mapk} in human tissues. A Northern blot containing 2 μg of poly(A)$^+$ mRNA isolated from different human tissues (He, heart muscle; Br, brain; Pl, placenta; Lu, lung; Li, liver; Sk, skeletal muscle; Ki, kidney; Pa, pancreas) was hybridized with probes for (A) p40^{mapk}/p41^{mapk} (nucleotides 787–1,458 of p41^{mapk} that correspond to nucleotides 701–1,372 of p40^{mapk}); (B) p44^{mapk} (nucleotides 7–191); (C) p63^{mapk} (nucleotides 1,673–2,058); and (D) actin. RNA size markers (in kb) are shown to the left of each panel.

kinase mRNAs suggests that different isoforms may play tissue-specific roles. Thus MAP kinases have been implicated in the M phase of the meiotic cycle [6,8,9], in the G_0-to-G_1 progression of quiescent cells [7,8,10], and may play a role in the differentiation of PC-12 cells [50,51]. Human tissues are physiologically regulated by a variety of extracellular factors that may employ a diversity of signal transduction pathways. It is possible that different isoforms of MAP kinases are expressed depending on the signal transduction needs of the differentiated cells present in each tissue.

Acknowledgements: We are grateful to Margaret Shepard for secretarial assistance and Debra A. Latour for technical assistance. This work was supported, in part, by Grants CA39240 and GM37845 from the National Institutes of Health. F.A.G. is a recipient of a postdoctoral fellowship from the National Science Foundation.

REFERENCES

[1] Ray, L.B. and Sturgill, T.W. (1987) Proc. Natl. Acad. Sci. USA 84, 1502–1506.
[2] Ray, L.B. and Sturgill, T.W. (1988) J. Biol. Chem. 263, 12721–12727.
[3] Rossomando, A.J., Payne, D.M., Weber, M.J. and Sturgill, T.W. (1989) Proc. Natl. Acad. Sci. USA 86, 6940–6943.
[4] Boulton, T.G., Yancopoulos, G.D., Gregory, J.S., Slaughter, C., Moomaw, C., Hsu, J. and Cobb, M.H. (1990) Science 249, 64–67.
[5] Ahn, N.G., Weiel, J.E., Chan, C.P. and Krebs, E.G. (1990) J. Biol. Chem. 265, 11487–11494.

[6] Sanghera, J.S., Paddon, H.B., Bader, S.A. and Pelech, S.L. (1990) J. Biol. Chem. 265, 52–57.
[7] Northwood, I.C., Gonzalez, F.A., Wartmann, M., Raden, D.L. and Davis, R.J. (1991) J. Biol. Chem. 266, 15266–15276.
[8] Posada, J., Sanghera, J., Pelech, S., Aebersold, R. and Cooper, J.A. (1991) Mol. Cell Biol. 11, 2517–2528.
[9] Rossomando, A.J., Sangheran, J.S., Marsden, L.A., Weber, M.J., Pelech, S.L. and Sturgill, T.W. (1991) J. Biol. Chem. 266, 20270–20275.
[10] Sturgill, T.W., Ray, L.B., Erikson, E. and Maller, J.L. (1988) Nature 334, 715–718.
[11] Gotoh, Y., Nishida, E., Matsuda, S., Shiina, N., Kosako, H., Shiokawa, K., Akiyama, T., Ohta, K. and Sakai, H. (1991) Nature 349, 251–254.
[12] Seth, A., Alvarez, E., Gupta, S. and Davis, R.J. (1991) J. Biol. Chem. 266, 23521–23524.
[13] Pulverer, B.J., Kyriakis, J.M., Avruch, J., Nikolakaki, E. and Woodgett, J.R. (1991) Nature 353, 670–674.
[14] L'Allemain, G., Pouyssegur, J. and Weber, M.J. (1991) Cell Reg. 2, 675–684.
[15] Ahn, N. and Krebs, E. (1990) J. Biol. Chem. 265, 11495–11591.
[16] Ahn, N.G., Seger, R., Bratlien, R.L., Diltz, C.D., Tonks, N.K. and Krebs, E.G. (1991) J. Biol. Chem. 266, 4220–4227.
[17] Gomez, N. and Cohen, P. (1991) Nature 353, 170–173.
[18] Campos-Gonzalez, R. and Glenney, J. (1991) Cell Reg. 2, 663–673.
[19] Ray, L.B. and Sturgill, T.W. (1988) Proc. Natl. Acad. Sci. USA 85, 3753–3757.
[20] Anderson, N.G., Maller, J.L., Tonks, N.K. and Sturgill, T.W. (1990) Nature 343, 651–653.
[21] Payne, D.M., Rossomando, A.J., Martino, P., Erickson, A.K., Her, J.-H., Shabanowitz, J., Hunt, D.F., Weber, M.J. and Sturgill, T.W. (1991) EMBO J. 10, 885–892.
[22] Boulton, T.G., Nye, S.H., Robbins, D.J., Ip, N.Y., Radziejewska, E., Morgenbesser, S.D., DePinho, N., Cobb, M.H. and Yancopoulos, G.D. (1991) Cell 65, 663–675.
[23] Seger, R., Ahn, N.G., Boulton, T.G., Yancopoulos, N.P., Radziejewska, L.E., Bratlein, R.L., Cobb, M.H. and Krebs, E.G. (1991) Proc. Natl. Acad. Sci. USA 88, 6142–6146.
[24] Crews, C.M., Alessandrini, A.A. and Erikson, R.L. (1991) Proc. Natl. Acad. Sci. USA 88, 8845–8849.
[25] Wu, J., Rossomando, A.J., Her, J.H., Del Vecchio, R., Weber, M.J. and Sturgill, T.W. (1991) Proc. Natl. Acad. Sci. USA 88, 9508–9512.
[26] Maller, J. (1991) Curr. Biol. 1, 334–335.
[27] Clark-Lewis, I., Sanghera, J.S. and Pelech, S.L. (1991) J. Biol. Chem. 266, 15180–15184.
[28] Gonzalez, F.A., Raden, D.L. and Davis, R.J. (1991) J. Biol. Chem. 266, 22159–22163.
[29] Erikson, A.K., Payne, D.M., Martino, P.A., Rossomando, A.J., Shabanowitz, J., Weber, M.J., Hunt, D.F. and Sturgill, T.W. (1990) J. Biol. Chem. 265, 19728–19735.
[30] Boulton, T.G., Gregory, J.S., Jong, S.M., Wang, L.H., Ellis, L. and Cobb, M.H. (1990) J. Biol. Chem. 265, 2713–2719.
[31] Bird, T.A., Sleath, P.R., deRoos, P.C., Dower, S.K. and Virca, G.D. (1991) J. Biol. Chem. 266, 22661–22670.
[32] Takishima, K., Griswold-Prenner, I., Ingebritsen, T. and Rosner, M.R. (1991) Proc. Natl. Acad. Sci. USA 88, 2520–2524.
[33] Alvarez, E., Northwood, I.C., Gonzalez, F.A., Latour, D.A., Seth, A., Abate, C., Curran, T. and Davis, R.J. (1991) J. Biol. Chem. 266, 15277–15285.
[34] Gotoh, Y., Moriyama, K., Matsuda, S., Okumura, E., Kishimoto, T., Kawasaki, H., Suzuki, K., Yahara, I., Sakai, H. and Nishida, E. (1991) EMBO J. 10, 2661–2668.
[35] Her, J.-H., Wu, J., Rall, T.B., Sturgill, T.W. and Weber, M.J. (1991) Nucleic Acids Res. 19, 3743.
[36] Kozak, M. (1987) Nucleic Acids Res. 15, 8125–8148.
[37] Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42–52.
[38] Courchesne, W.E., Kunisawa, R. and Thorner, J. (1989) Cell 58, 1107–1119.
[39] Elion, E.A., Grisafi, P.L. and Gink, G.R. (1990) Cell 60, 649–664.
[40] Elledge, S.J. and Spottswood, M.R. (1991) EMBO J. 10, 2653–2659.
[41] Tsai, L.-H., Harlow, E. and Meyerson, M. (1991) Nature 353, 174–177.
[42] Ninomiya-Tsuji, J., Nomoto, S., Yasuda, H., Reed, S.I. and Matsumoto, K. (1991) Proc. Natl. Acad. Sci. USA 88, 9006–9010.
[43] Lee, M.G. and Nurse, P. (1987) Nature 327, 31–35.
[44] Rich, B.E. and Steitz, J.A. (1987) Mol. Cell Biol. 7, 4065–4074.
[45] Pari, G., Jardine, K. and McBurney, M.W. (1991) Mol. Cell Biol. 11, 4796–4803.
[46] Ben-Neriah, Y., Bernards, A., Paskind, M., Daley, G.Q. and Baltimore, D. (1986) Cell 44, 577–586.
[47] Grove, J.R., Banerjee, P., Blasubramanyam, A., Coffer, P.J., Price, D.J., Avruch, J. and Woodgett, J.R. (1991) Mol. Cell Biol. 11, 5541–5550.
[48] Hanks, S.K. (1991) Curr. Opinion Struct. Biol. 1, 369–383.
[49] Kozak, M. (1991) J. Cell Biol. 115, 887–903.
[50] Gomez, N., Tonks, N.K., Morrison, C., Harmar, T. and Cohen, P. (1990) FEBS Lett. 271, 119–122.
[51] Miyasaka, T., Chao, M.V., Sherline, P. and Saltiel, A.R. (1990) J. Biol. Chem. 265, 4730–4735.