Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 173–177, January 1993
Biochemistry

Molecular structure of a protein-tyrosine/threonine kinase activating p42 mitogen-activated protein (MAP) kinase: MAP kinase kinase

(molecular cloning/byr1/STE7)

JIE WU*†, JEFFREY K. HARRISON†, LEIGH ANN VINCENT*†, CLAIRE HAYSTEAD†, TIMOTHY A. J. HAYSTEAD†, HANSPETER MICHEL‡, DONALD F. HUNT‡, KEVIN R. LYNCH†, AND THOMAS W. STURGILL*†§

Departments of *Internal Medicine, †Pharmacology, and ‡Chemistry, University of Virginia, Charlottesville, VA 22908

Communicated by Oscar L. Miller, Jr., October 9, 1992

ABSTRACT **MAP kinases $p42^{mapk}$ and $p44^{mapk}$ participate in a protein kinase cascade(s) important for signaling in many cell types and contexts. Both MAP kinases are activated *in vitro* by MAP kinase kinase, a protein-tyrosine and threonine kinase. A MAP kinase kinase cDNA was isolated from a rat kidney library by using peptide sequence data we obtained from MAP kinase kinase isolated from rabbit skeletal muscle. The deduced sequence, containing 393 amino acids (predicted mass, 43.5 kDa), is most similar to byr1 (Bypass of ras1), a yeast protein kinase functioning in the mating pathway induced by pheromones in *Schizosaccharomyces pombe*. An unusually large insert is present in MAP kinase kinase between domains IX and X and may contribute to protein–protein interactions with MAP kinase. Major (2.7 kilobases) and minor (1.7 kilobases) transcripts are widely expressed in rat tissues and appear to be derived from a single gene.**

Members of a family of serine/threonine kinases that require phosphorylation of both tyrosine and serine/threonine residues for activation in response to external stimuli, known collectively as MAP kinases, have been shown to participate in protein kinase cascades in yeast, amphibians, and mammals (for review, see ref. 1). Included are the mammalian proteins $p42^{mapk}$, $p44^{mapk}$ (2, 3), and p55 MAP kinase, a distinct enzyme whose sequence has not yet been reported (4). In addition, several yeast MAP kinases have been identified, including FUS3 (5) and KSS1 (6) in the pheromone-response pathway for mating in *Saccharomyces cerevisiae* (7) and spk1, the probable equivalent of KSS1/FUS3 in *Schizosaccharomyces pombe* (8). A requirement for dual phosphorylation on tyrosine and threonine residues has been demonstrated for p42 and p44 MAP kinases (1), p55 MAP kinase (4), and FUS3 (7).

Ahn *et al.* (9) discovered an activator of p42 and p44 MAP kinases that we (ref. 10 and references therein) and others have identified as a protein-tyrosine and threonine kinase, termed MAP kinase kinase. MAP kinase kinase is remarkably specific for p42/p44 MAP kinases as substrates; no exogenous protein or peptide substrate has been identified for MAP kinase kinase other than p42/p44 MAP kinases (11). The latter proteins are phosphorylated by MAP kinase kinase on the regulatory threonine and tyrosine residues present in a TEY amino acid motif just upstream of the conserved APE amino acid sequence (1).

To identify the MAP kinase kinase phosphorylating and activating $p42^{mapk}$ for molecular cloning, we partially purified MAP kinase kinase from rabbit skeletal muscle and identified protein bands from SDS/polyacrylamide gel electrophoresis for sequencing by renaturation of enzyme activity (12). Activity was found in two bands of 46 and 48 kDa and the major proteins therein identified as a kinase and a protein related to smgp25 GDP dissociation inhibitor, respectively. We now report the results of molecular cloning of a cDNA of MAP kinase kinase.¶ Analysis of the deduced amino acid sequence reveals sequence similarity to yeast kinases thought to be the immediate upstream activators of MAP kinases in the mating pathways of budding and fission yeast.

MATERIALS AND METHODS

Materials. Oligonucleotides, the Sequenase II kit, *Taq* DNA polymerase, and nylon transfer membrane were purchased from Operon Technologies (Alameda, CA), United States Biochemical, Promega, and MSI (Westboro, MA), respectively. Restriction enzymes, random-primer labeling kit, and oligo(dT)-cellulose were from GIBCO/BRL.

PCR and cDNA Cloning. PCR primer 7-1 (sense; 5′-GTIWSICAYAARCCIWSIGGIYTIGTIATGGCIMG), primer 7-2 (antisense; 5′-RTTGACCAGGATRTTGSWCGGYTTIACRTC), and primer 7-3 (sense; 5′-GARCTGGGCGCCGGCAWCGGCGGCGTGGTGTYAAR) (R, purine; W, A or T; Y, pyrimidine; S, G or C; M, A or C) were synthesized based on partial sequence (underlined) of peptides VSHKPSGLVMAR [peptide 7-1 (12)], DVKPSNILVNSR [peptide 7-2 (12)], and ISELGAGN(I)GGVVFK (peptide 7-3) obtained and sequenced as described (12).

Double-stranded cDNAs were synthesized from poly(A)$^+$ RNA by oligo(dT) priming. The PCR was performed for 30 cycles of 94°C for 1.5 min, 40°C for 2.5 min, and 72°C for 3.5 min. PCR products were diluted 1:500 and reamplified for 30 cycles of 94°C for 1.5 min, 55°C for 2.5 min, and 72°C for 3.5 min.

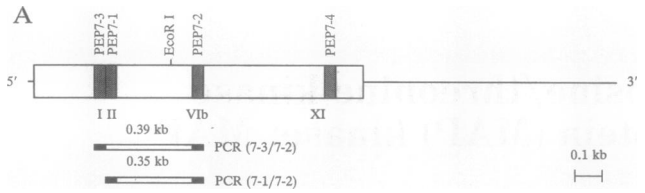
The 0.35-kilobase (kb) PCR product from rat kidney cDNA obtained by using primer pair 7-1/7-2 was gel-purified and ^{32}P-labeled by the random-primer method for use as a probe to screen a rat kidney λgt10 cDNA library (13). Hybridizations were carried out at 65°C in 6× SSC (1× SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7)/5× Denhardt's solution (1× Denhardt's solution is 0.02% Ficoll 400/0.02% bovine serum albumin/0.02% polyvinylpyrrolidone)/0.1% SDS/2 mM EDTA containing denatured salmon sperm DNA (0.1 mg/ml). Filters were washed four times (30 min, 22°C) with 1× SSC/0.1% SDS and two times (20 min, 65°C) with 0.2× SSC/0.1% SDS. Approximately 10^6 plaques were screened. Inserts from positive plaques were amplified by PCR (λgt10

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: MAP, mitogen-activated protein.

§To whom reprint requests should be addressed at: Box 419, Department of Internal Medicine, University of Virginia Health Sciences Center, Charlottesville, VA 22908.

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. L04485).



B

```
  1: gaattccgcggaattccgCCCGGCTGCAAGATGCCCAAGAAGAAGCCGACGCCCATCCAG
  1:                                M  P  K  K  K  P  T  P  I  Q

 61: CTGAACCCGGCCCCCGATGGCTCCGCGGTTAACGGGACCAGCTCGGCCGAGACCAACCTG
 11:  L  N  P  A  P  D  G  S  A  V  N  G  T  S  S  A  E  T  N  L

121: GAGGCCTTGCAGAAGAAGCTGGAGGAGCTGGAGCTGGACGAGCAGCAGCGGAAGCGCCTT
 31:  E  A  L  Q  K  K  L  E  E  L  E  L  D  E  Q  Q  R  K  R  L

181: GAGGCCTTTCTGACGCAGAAGCAGAAGGTGGGAGAGTTGAAGGATGATGACTTTGAGAAG
 51:  E  A  F  L  T  Q  K  Q  K  V  G  E  L  K  D  D  D  F  E  K

241: ATCAGTGAACTGGGGGCTGGCAATGGTGGAGTGGTGTTCAAGGTCTCCCACAAGCCATCT
 71:  I  S  E  L  G  A  G  N  G  G  V  V  F  K  V  S  H  K  P  S
                  Peptide 7-3                    Peptide 7-1
301: GGCCTGGTTATGGCTAGGAAGCTAATTCACCTGGAGATCAAACCCGCAATCCGGAACCAG
 91:  G  L  V  M  A  R  K  L  I  H  L  E  I  K  P  A  I  R  N  Q

361: ATCATCCGGGAGCTGCAGGTGCTGCATGAGTGCAACTCCCCGTACATAGTGGGCTTCTAC
111:  I  I  R  E  L  Q  V  L  H  E  C  N  S  P  Y  I  V  G  F  Y

421: GGGGCCTTCTACAGTGACGGCGAGATCAGCATCTGCATGGAGCACATGGATGGTGGGTCC
131:  G  A  F  Y  S  D  G  E  I  S  I  C  M  E  H  M  D  G  G  S

481: TTGGATCAAGTGCTGAAGAAAGCTGGAAGAATTCCTGAGCAAATTTTAGGAAAAGTCAGC
151:  L  D  Q  V  L  K  K  A  G  R  I  P  E  Q  I  L  G  K  V  S

541: ATCGCTGTGATAAAAGGCCTGACATATCTACGAGAGAAGCACAAGATTATGCACAGAGAT
171:  I  A  V  I  K  G  L  T  Y  L  R  E  K  H  K  I  M  H  R  D

601: GTCAAGCCTTCCAACATTCTAGTGAACTCACGTGGGGAGATCAAACTCTGCGATTTTGGG
191:  V  K  P  S  N  I  L  V  N  S  R  G  E  I  K  L  C  D  F  G
                  Peptide 7-2
661: GTCAGCGGGCAGCTAATTGACTCCATGGCCAACTCCTTCGTGGGAACAAGGTCCTACATG
211:  V  S  G  Q  L  I  D  S  M  A  N  S  F  V  G  T  R  S  Y  M

721: TCGCCTGAGAGACTCCAGGGGACTCACTACTCTGTGCAGTCGGACATCTGGAGCATGGGG
231:  S  P  E  R  L  Q  G  T  H  Y  S  V  Q  S  D  I  W  S  M  G

781: CTCTCTCTGGTGGAGATGGCAGTTGGAAGATACCCCATTCCTCCTCCTGATGCCAAGGAG
251:  L  S  L  V  E  M  A  V  G  R  Y  P  I  P  P  P  D  A  K  E

841: CTGGAGCTGCTGTTTGGATGCCAGGTGGAAGGAGACGCGGCCGAAACGCCACCCAGGCCA
271:  L  E  L  L  F  G  C  Q  V  E  G  D  A  A  E  T  P  P  R  P

901: AGGACCCCTGGGAGGCCCCTCAGCTCATATGGAATGGATAGCCGACCTCCCATGGCAATT
291:  R  T  P  G  R  P  L  S  S  Y  G  M  D  S  R  P  P  M  A  I

961: TTTGAGTTGTTGGATTACATCGTCAATGAGCCTCCTCCAAAACTGCCCAGTGGAGTATTC
311:  F  E  L  L  D  Y  I  V  N  E  P  P  P  K  L  P  S  G  V  F

1021: AGTCTGGAATTTCAGGATTTTGTGAATAAGTGCTTAATAAAGAACCCTGCAGAGAGAGCA
331:  S  L  E  F  Q  D  F  V  N  K  C  L  I  K  N  P  A  E  R  A

1081: GATCTGAAGCAGCTCATGGTACATGCTTTCATCAAGAGATCTGATGCCGAGGAGGTAGAC
351:  D  L  K  Q  L  M  V  H  A  F  I  K  R  S  D  A  E  E  V  D
                        Peptide 7-4
1141: TTCGCAGGCTGGCTCTGCTCCACCATTGGGCTTAACCAGCCCAGCACACCAACCCACGCT
371:  F  A  G  W  L  C  S  T  I  G  L  N  Q  P  S  T  P  T  H  A

1201: GCCAGCATCTGAGCCTTTGGGAAGCAGCAGAGAGGAATCCTCTGCCCAGTGGCATGCCAT
391:  A  S  I  @

1261: GCCGCTTTCAGGCCTCTCCCATGCTTGTCTATGTTCAGATGTGCATCTCATCTGTGACAA
1321: AGGATGAAGAACACAGCATGTGCCAAATCGTACTTGTGTCATTTTTAATATTGTCTTTAT
1381: CGCTATGGTTACTCCCCTAAGTGGATTGGCTTTGTGCTTGGGGCTATTTGTCTGTTCATC
1441: AAATACATGCCAGGTTGAACTACAGTGAAACCCTGGTGACCTGGGTGGTCTTCTTACTGA
1501: TGTTTGCGCTGCTGTTCATCGTGACTCACTAGCTGGCTGCCTGTATTGTCAGGATTCTCG
1561: GACCCTTGGTACTTCACTCTTGCTGGTGACCTCTCAGTCTGAGGAGAGGGGGCCTTCTGA
1621: GACCCTTCACAGGCAGTGCATGCATGAAAAGCATGCTTTGCTGCTACTGAAATGAGCACC
1681: AGAACGTGTACATCATGGTATTTTATTTTTGCTTTTGGTATAGAACTCAGCAGTTCCCAT
1741: TTAAAAAAAAAAATCTAACCAGAGCCCATCACTGCCATGATAGCTGGGGCTTCAGTCTGTC
1801: TACTGTGGTGATTTTAGACTTCTGGTTGTATTTCTATATTTATTTTTAAATATACTGTGT
1861: GGGATACTTAGTGGTATATGTCTCTGAGTTTGGATTAGTGTTTCTAAATTGGTAGTTATT
1921: TTGAATGTCACAAATGGATTAAGGAATCAACGTATCAAGAGTTCTATCTTTCTTCCAGTC
1981: TAAGTACCAATGCTATTGTAAACGTGTATAGTGCCTACAAATTGTATGAAAACCCTTTTA
2041: ACCACTTTACTCAAGATGTTTATCAAATCTAATCTCTTATTCTAATAAAAATACTATCAA
2101: GTTAAAGTAAAAAAAAAAAAAAAAAAAAcggaattc
```

FIG. 1. MAP kinase kinase cDNA from rat kidney. (*A*) Schematic representation of pK28 depicting location of protein sequences (shaded) that match tryptic peptide sequences from rabbit skeletal muscle and the region between domains I/II and VIb used to generate a probe by PCR. Open box, coding region; solid line, 5′ (left) and 3′ (right) untranslated regions. Strategy for PCR and primers (solid bars) are shown below the map. (*B*) Nucleotide sequence of pK28 and deduced amino acid sequence of MAP kinase kinase in the single-letter code. Lowercase type identifies nucleotide sequences derived from *Eco*RI linkers used in generating the library (13). Boldface type, translational start codon and the polyadenylylation signals; @, translational stop codon; double underlines and underlines, amino acid sequences that match peptide sequences from MAP kinase kinase purified from rabbit (this report and refs. 12 and 16) and mouse (17), respectively. Where mouse peptide sequences overlap with rabbit peptide sequence, only rabbit peptide sequence is indicated.

forward and reverse primers) and analyzed by agarose gel electrophoresis and Southern blot hybridization. The insert from clone K28 (see text) was subcloned into pBluescript SK− (Stratagene) after partial digestion with *Eco*RI, generating plasmid pK28. DNA sequencing of pK28 was performed by the dideoxynucleotide chain-termination method using a Sequenase II kit. Both strands of DNA were sequenced.

Sequence Analysis. Protein sequence data bases, [National Biomedical Research Foundation (Release 33), GenPept (Release 72), and Swiss-Prot (Release 22)] were searched by the FASTA program (14). Programs BESTFIT and PIPEUP (GCG, Madison, WI) were used to compare and align MAP kinase kinase-related protein sequences.

Other Methods. Hybridization with oligonucleotide probe 7-1 was carried out overnight at 32°C in 6× SSC/5× Denhardt's solution/50 mM sodium phosphate, pH 7.0/2 mM EDTA/0.1% SDS containing denatured salmon sperm DNA (0.05 mg/ml) and ^{32}P-labeled oligonucleotide (1 pmol/ml). The filter was washed for three 30-min periods at room temperature with 2× SSC/0.1% SDS and for two 30-min periods at 37°C with 1× SSC/0.1% SDS.

Northern and Southern blot analyses were carried out essentially as described (15). The probe (1.6 kb; nucleotides 282–1892) was produced by PCR from the MAP kinase kinase cDNA and labeled with ^{32}P by using the random-primer method.

RESULTS

Peptide Sequencing of MAP Kinase Kinase. We obtained sequences of four peptides from MAP kinase kinase, purified from rabbit skeletal muscle (12) (Fig. 1). Peptides 7-1 and 7-2 are from conserved kinase subdomains II and VIb, respectively, and have been reported (12). Peptide 7-3 is from conserved subdomain I and contains the conserved GXGXXG motif present in the ATP-binding site. Peptide 7-4 [QLMVHAFI(N)K] does not correspond to any conserved kinase subdomain. Peptides 7-3 and 7-1 were expected to be contiguous, based on alignment with conserved sequences of protein kinases.

Amplification of a Partial cDNA by PCR. Oligonucleotide primers for amplification by PCR were synthesized based on the sequence of peptides 7-1, 7-2, and 7-3 and their orientation, predicted by comparison with known protein kinases (Fig. 1*A*). Amplification of double-stranded cDNA synthesized from poly(A)$^+$ RNA prepared from rat brain, tongue, kidney, or liver by using primers 7-3 and 7-2, yielded a DNA fragment of 0.39 kb in each case (data not shown). The size of the fragment is approximately the size expected from spacing of amino acids between domains I and VIb of protein kinases (data not shown). To confirm that the PCR product was the correct one, each amplified DNA was shown to hybridize to an oligonucleotide (primer 7-1) derived from domain II (data not shown). Amplification of the 0.39-kb fragment from each tissue by using the internal primer 7-1 and primer 7-2 yielded the expected 0.35-kb fragment. The 0.39-kb fragments from tongue and kidney were subcloned. Two clones from each tissue were sequenced and found to be identical in the regions examined and to encode peptide 7-1 (data not shown). Therefore, each tested tissue appeared to contain a similar mRNA. The 0.35-kb DNA fragment from kidney cDNA was used as a probe for library screening.

Isolation of a cDNA for MAP Kinase Kinase. Thirty-one positive clones were obtained by screening a cDNA library

from rat kidney and purified. cDNA inserts from 29 of the 31 positive plaques were amplified by PCR and analyzed by Southern blot hybridization with oligonucleotide 7-1. Ten of the 29 clones analyzed contained cDNA inserts that hybridized with the oligonucleotide probe (data not shown). Clone K28, containing the largest insert among these 10 clones, was subcloned to generate plasmid pK28 and sequenced. Plasmid pK28 contained a 2136-base-pair insert with an open reading frame from nucleotides 31 to 1209. This open reading frame encodes a protein of 393 amino acid residues with a predicted molecular mass of 43.5 kDa. A polyadenylylation signal, AATAAA, is located 20 nucleotides upstream of the polyadenylylation site (Fig. 1*B*).

The deduced protein sequence from clone pK28 contains amino acid sequences that match exactly the sequences determined for four tryptic peptides from rabbit skeletal muscle (peptides 7-1 to 7-4). The start codon of the open reading frame is not in a sequence context predicted as optimal for a Kozak translation initiation sequence (18). However, this ATG likely encodes the initiation methionine because it is followed by an amino acid sequence that matches 15 of 15 amino acid residues sequenced from the N terminus of MAP kinase kinase from rabbit skeletal muscle (16) (Fig. 1*B*) and 14 of 16 residues from the N terminus of MAP kinase kinase from *Xenopus* eggs (19). Therefore, the DNA insert of pK28 encodes a full-length MAP kinase kinase from rat kidney. The initiation methionine appears to be processed after translation (Fig. 1*B*).

Clones hybridizing to the 0.35-kb probe but not to the internal oligonucleotide 7-1 probe have not been characterized. Hybridization to the 0.35-kb probe indicates that these cDNAs have related sequences; failure to hybridize to the oligonucleotide probe suggests that these clones are not identical to MAP kinase kinase. Thus, they may encode protein kinases related to MAP kinase kinase.

Primary Structure of MAP Kinase Kinase. The deduced protein sequence contains all of the conserved motifs identified by Hanks and Quinn (20) for protein kinases (Fig. 2). No definite diagnostic motifs exist to separate traditional protein kinases from "dual specificity kinases" that phosphorylate both tyrosine and serine/threonine residues (25). Threonine is present at residue 226, and serine or threonine in this position is strongly correlated with serine/threonine specificity (25). A small noncharged amino acid (serine, underlined) is present between proline and asparagine in DVKPSN in domain VIB and this loosely correlates with dual specificity (12, 25).

Searches of current protein sequence data bases using the FASTA algorithm (14) revealed that the sequence of rat MAP kinase kinase was distinct from any sequence therein but was found to be most similar to four yeast protein kinases: byr1 (21), wis1 (22), PBS2 (23), and STE7 (24), with optimal scores of 638, 585, 571, and 375, respectively. In contrast, scores for other protein kinases were in the range of 200–300 (data not shown). MAP kinase kinase is most similar in size to byr1 (393 versus 340 amino acids, respectively). Values for identity (similarity) of aligned sequences, provided by program BESTFIT, were 45% (65%), 42% (61%), 42% (60%), and 39% (61%) for comparison of MAP kinase kinase to byr1, wis1, PBS2, and STE7, respectively.

MAP kinase kinase has a proline-rich segment of 46 amino acids between domains IX and X (residues 262–307), containing 11 of the 29 proline residues in the protein. The 11 proline residues are clustered at the beginning and end of the segment. Residues between domains IX and X in cAMP-dependent protein kinase form a loop connecting helices F and G present in the crystal structure of the catalytic subunit and the loop lies below the enzymatic cleft (26). The "normal" spacing between helices F and G exhibited by the majority of kinases is 10 or 11 residues (20, 26). A large "gap" is required for alignment of each yeast kinase with MAP kinase kinase (Fig. 2), indicating that MAP kinase kinase contains an insert in this region. A few kinases in the database have short inserts (20, 26). However, the looping segment in MAP kinase kinase contains 49 residues. This may be an important feature for determining specific recognition of protein substrate.

The N-terminal sequence (PKKKPTPI) resembles the nuclear localization signal of simian virus 40 large tumor antigen (27). In addition, a bipartite motif (28) correlating with transport of proteins into the nucleus is located between Lys-34 and Arg-49. However, the functional significance of these sequences is as yet undetermined.

Northern and Southern Blot Analysis. To assess the extent of mRNA expression and the number and size of the transcripts, poly(A)$^+$ RNA was isolated from rat tissues and subjected to Northern blot analysis (Fig. 3). MAP kinase kinase mRNA was expressed in each tissue examined, consistent with the results of PCR amplification. The probe detected a major 2.7-kb transcript and a minor 1.7-kb transcript in each adult tissue. An additional 2.3-kb transcript was present in neonatal rat tongue, possibly reflecting developmental regulation. All three messages are sufficiently large to encode a full-length MAP kinase kinase. Based on correspondence in size, the sequenced clone (K28) is most likely derived from the 2.7-kb message. The other mRNAs may encode the same protein or a closely related protein.

Within the brain (Fig. 3*A*), expression of mRNA was most prominent in olfactory bulb, hippocampus, and basal ganglia. Expression was fairly uniform in other regions examined. Outside the brain (Fig. 3*B*), mRNAs for MAP kinase kinase appeared to be most abundant in spinal cord, tongue, large intestine, aorta, and diaphragm. Intestine and aorta contain

```
MKK   MPKKKPTPIQLNPAPDGSAVNGTSSAETNLEALQKKLEELELDEQQRKRLEAFLT-----   55
byr1  .F.RRRN.KG.VLN.NA.VKSSDNDHKEE.INN..SF.---------SNV...ME-----   46
wis1  IRRAP.GKLD.SNSNPT.PVSPS.M.SRRGLNIPPT.KQ-AVS.TPFSTFSDI.DAKSGT  291
PBS2  STEGPHDTVGTT.RTGN.NNSSN.GSS--------------GGGGLFANFSKYVDIKSGS  331
STE7  -SSSLS..CIIDAYSNNFGLSPS.TNS.PSTIQGLSNIATPVENEHSIS.PPLEESLSPA  169

                                         * *  * *            * *
MKK   ---------Q-------KQKVGELKDDDFEKISELGAGNGGVVFKVSHKPSGLVMARKLI   99
byr1  ---------.CAHMNRRPAWISD.DNSSL.VVRH..E....A.SL.K.R--NIF....TV   95
wis1  LNFKNKAVLNSEGVNFSSGSSFRINMSEII.LE...K..Y...Y.AL.Q.T.VT..L.E.  351
PBS2  LNFAGKLSLSSKGIDFSNGSSSRITL.EL.FLD...H..Y.N.S..L...TNVI..T.EV  391
STE7  AADLKDTLSGTSNGNY-------IQLQ.LVQLGKI....S.T.V.AL.V.DSKIV.K.T.  222
                                         I                   II

                     *            *
MKK   HLEI-KPAIRNQIIRELQVLHECNSP-YIVGFYGAFY---SDGEISICMEHMDGGSLDQV  154
byr1  YVGS-DSKLQK..L...G...H.R..-.........Q---YKNN..L...Y..C....AI  150
wis1  R.SL-EE.TF...IM..DI..KAV..-.........F---VE.SVF....Y..A..M.KL  406
PBS2  R..L-DE.KFR...M..E...K....-.........F---IE.AVYM...Y.......KI  446
STE7  PVEQNNST.I..LV...SIVKNVKPHEN.IT....Y.NQHINNE.I.L..YS.C....KI  282
                     III          IV                 V

                                                      * *  *
MKK   LK----------KA---GRIPEQILGKVSIAVIKGLTYLREKHKIMHRDVKPSNILVNS-  200
byr1  .R----------EG---.P..LD....IINSMV...I..YNVLH.I...L....VV...-  196
wis1  YA----------G----.IKD.GV.ARTAY..VQ..KT.K.E.N.I......T.V....N  452
PBS2  YD----------ESSEI.G.D.PQ.AFIAN...H..KE.K.Q.N.I......T...CSAN  496
STE7  .SVYKRFVQRGTVSSKKTWFN.LTIS.IAYG.LN..DH.YRQY..I...I....V.I..-  341
                                            VIa         VIb

           *  **                     **                    * ** *
MKK   RGEIKLCDFGVSGQLIDSMANSFVGTRSYMSPERLQGTH-------YSVQSDIWSMGLSL  253
byr1  ............E.VN.V.QT....ST......IR.GK-------.T.K.....L.I.I  249
wis1  -.QV.........N.VA.ISKTNI.CQ...A...IRVGGPTNGVLT.T..A.V..L..TI  511
PBS2  Q.TV.........N.VA.L.KTNI.CQ...A...IKSLNPDRA--T.T.......L...I  554
STE7  K.Q.........KK..N.I.DT....ST......I..N-------V..IKG.V..L..MI  394
              VII                    VIII                  IX

MKK   VEMAVGRYPIPPPDAKELELLFGCQVEGDAAETPPRPRTPGRPLSSYGMDSRPPMAIFEL  313
byr1  I.L.TQEL.------------------------------WSF.NID----DSIG.LD.  273
wis1  L...L.A..Y..ESYT--------------------------------------S..AQ  532
PBS2  L...L....Y..ETYD--------------------------------------N..SQ  575
STE7  I.LVT.EF.LGGHNDT-------------------------------------.DG.LD.  417

                                          *
MKK   LDYIVNEPPPKLPSG-VFSLEFQDFVNKCLIKNPAERADLKQLMVHAFIKRS--DAEEVD  370
byr1  .HC..Q.E..R...S--.PEDLRL..DA..H.D.TL..SPQ..CAMPYFQQA--LMIN..  329
wis1  .SA.CDGD..S..-D-S..PEAR.......N...SL.P.YHE.AN.PWLLKY--QNAD..  588
PBS2  .SA..DG...R...D-K..SDA....SL..Q.I.ER.PTYAA.TE.PWLVKY--RNQD.H  632
STE7  .QR.....S.R..KDRIY.K.MT....R.C...ERE.SSIHE.LH.DL.MKYVSPSKDDK  477
       X                                  XI

MKK   FAGWLCSTIGLNQPSTPTHAASI@                                      393
byr1  L.S.ASNFRSS@                                                  340
wis1  M.S.AKG-------ALKEKGEKRS@                                     605
PBS2  MSEYITERLERRNKILRERGENGLSKNVPALHMGGYSVNIQIKANRHVNITTKKKQTKSY  692
STE7  FRH.CRKIKSKIKEDKRIKREALDRAKLEKKQSERSTH@                       515
```

FIG. 2. Alignment of amino acid sequences of MAP kinase kinase (MKK) with byr1 (21), wis1 (22), PBS2 (23), and STE7 (24). Dots, identical amino acids; dashes, gaps introduced for alignment; @, C termini; roman numerals, protein kinase domains; *, conserved amino acid residues (20).

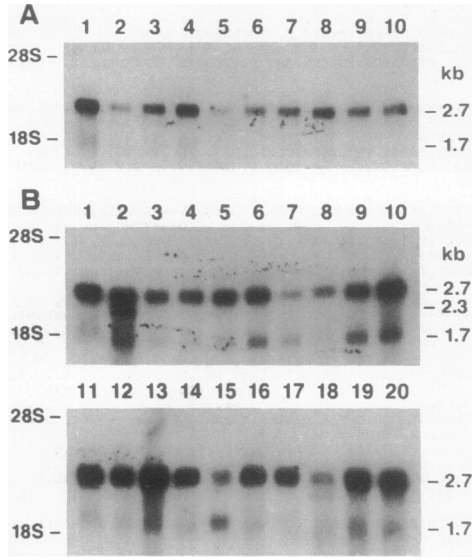


FIG. 3. Expression of MAP kinase kinase mRNA in rat tissues. mRNA for MAP kinase kinase was detected in adult tissues within the brain [3 μg of poly(A)$^+$ RNA per lane] (*A*) and tissues outside the brain [5 μg of poly(A)$^+$ RNA per lane] (*B*) by Northern blot analyses. Tongue was neonatal tissue. (*A*) Lanes: 1, olfactory bulb; 2, cortex; 3, hippocampus; 4, basal ganglia; 5, thalamus; 6, hypothalamus; 7, colliculi; 8, midbrain; 9, pons; 10, cerebellum. (*B*) Lanes: 1, spinal cord; 2, neonatal tongue; 3, atria; 4, ventricle; 5, lung; 6, kidney; 7, liver; 8, stomach; 9, small intestine; 10, large intestine; 11, aorta; 12, adrenal; 13, diaphragm; 14, bladder; 15, uterus; 16, testis; 17, seminal vesicle; 18, prostate; 19, epididymis; 20, vas deferens. Blots were processed concurrently.

smooth muscle whereas tongue and diaphragm are predominantly skeletal muscle.

Hybridization of MAP kinase kinase cDNA to rat genomic DNA detected a single 8.2-kb *Hin*dIII fragment. Two hybridizing fragments (6.2 kb and 5.3 kb) were detected in *Eco*RI-digested DNA (Fig. 4). However, the cDNA contains an *Eco*RI site (Fig. 1*A*). Thus, it appears that there is a single MAP kinase kinase gene. If true, the three transcripts observed by Northern blot analyses may result from differences in transcription or processing.

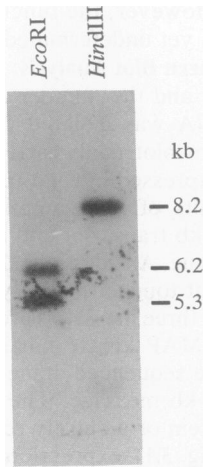


FIG. 4. Southern blot of rat genomic DNA. DNA from rat liver was digested with the indicated restriction enzyme, fractionated (15 μg) by 0.9% agarose gel electrophoresis, and hybridized to the ^{32}P-labeled 1.6-kb MAP kinase kinase fragment. The size of hybridizing fragments is indicated to the right.

DISCUSSION

We have isolated and characterized a cDNA for MAP kinase kinase from rat kidney. The predicted molecular mass agrees well with the apparent mass on SDS/polyacrylamide gels for homogenous MAP kinase kinase prepared from *Xenopus* eggs (19), A431 cells (11), skeletal muscle (16), and murine T cells (17) and the mass of renatured band 7 (12). The predicted sequence contains all four peptides we sequenced from MAP kinase kinase from rabbit skeletal muscle (ref. 12 and this report). Because these peptides were sequenced from an SDS/polyacrylamide gel band proven to have MAP kinase kinase enzymatic activity by renaturation (12), this correspondence indicates that the protein we have cloned is the MAP kinase activator. The sequence also contains exact matches to N-terminal peptide sequence from nearly homogeneous MAP kinase kinase from rabbit (16) and four internal peptides from the homogeneous mouse enzyme (17), suggesting that MAP kinase kinase may be highly conserved between species. Furthermore, transient expression of the cDNA in COS cells amplifies serum-stimulated MAP kinase kinase activity (J.W. and C.H., unpublished data). Thus these data establish the authenticity of the cloned cDNA.

The protein sequence of MAP kinase kinase had the highest similarity scores to four yeast protein kinases: byr1 (21) and wis1 (22) from *Schizosaccharomyces pombe* and PBS2 (23) and STE7 (24) from *Saccharomyces cerevisiae*. Most notable is the similarity to byr1 and STE7. byr1 and STE7 function in pathways for mating and conjugation/sporulation initiated by pheromones in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*, respectively. Pertinent here is that both pathways contain kinases [spk1 (8) and FUS3/KSS1 (5, 6) with sequence similarity to MAP kinases; spk1, FUS3, and KSS1 all contain sequences corresponding to the regulatory phosphorylation sites found in p42/p44 MAP kinases]. Phosphorylation of tyrosine and threonine in the TEY motif is required for FUS3 function (7) and, presumptively, required for spk1 and KSS1 functions. spk1 and FUS3/KSS1 are thought to be the immediate downstream targets of byr1 and STE7, respectively (7). Thus, it is reasonable to expect the MAP kinase activator to have sequence similarity to byr1 and STE7.

All of the "dual specificity kinases" that have been described autophosphorylate tyrosine but phosphorylate serine/threonine residues in exogenous substrates and in that sense are considered dual specificity kinases. The only protein kinases that have been shown to have dual specificity with exogenous substrates are MAP kinase kinase and (presumptively) members of this family and wee1 (25), all of which are kinase kinases. The fact that no sequence motifs have been defined that are diagnostic for dual specificity may reflect this.

Although the yeast mating pathways do not have a direct parallel to MAP kinase pathways in higher eukaryotes, their functions are similar when viewed as pathways for extracellular control of gene transcription. However, the pathways are clearly distinct, even between budding and fission yeast. Expression of p42^{mapk} does not complement FUS3/KSS1, implying that it is not a substrate for STE7 or, alternatively, that p42^{mapk} does not have substrate specificity equivalent to FUS3/KSS1 (G. Ammerer, personal communication). It is likely that different MAP kinases will have specific activator proteins.

We thank P. Dent and M. Weber for discussions and Tom Rall for technical assistance. Primary support was provided by grants to T.W.S. from the National Institutes of Health (DK41077) and American Cancer Society (BE-69D). Fellowship support for J.W. and J.K.H. was provided by the Juvenile Diabetes Foundation, Intl. and a National Research Service Award Fellowship HL08223, respectively. Ancillary support was provided by National Institutes of

Health grants to D.F.H. (GM37537), to K.R.L. (PO1 HL19242), and the University of Virginia Diabetes Center. T.A.J.H. is supported by the Markey Foundation.

1. Pelech, S. & Sanghera, J. S. (1992) *Science* **257**, 1355–1356.
2. Boulton, T. G., Yancopoulos, G. D., Gregory, J. S., Slaughter, C., Moomaw, C., Hsu, J. & Cobb, M. (1990) *Science* **249**, 64–67.
3. Her, J.-H., Wu, J., Rall, T. B., Sturgill, T. W. & Weber, M. J. (1991) *Nucleic Acids Res.* **19**, 3743.
4. Kyriakis, J. M., Brautigan, D. L., Ingebritsen, T. S. & Avruch, J. (1991) *J. Biol. Chem.* **266**, 10043–10046.
5. Elion, E. A., Grisafi, P. L. & Fink, G. R. (1990) *Cell* **60**, 649–664.
6. Courchesne, W. E., Kunisawa, R. & Thorner, J. (1992) *Cell* **58**, 1107–1119.
7. Gartner, A., Nasmyth, K. & Ammerer, G. (1992) *Genes Dev.* **6**, 1280–1292.
8. Toda, T., Shimanuki, M. & Yanagida, M. (1991) *Genes Dev.* **5**, 60–73.
9. Ahn, N. G., Seger, R. L., Bratlein, R. L., Diltz, C. D., Tonks, N. K. & Krebs, E. G. (1991) *J. Biol. Chem.* **265**, 11487–11494.
10. Rossomando, A., Wu, J., Weber, M. & Sturgill, T. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5221–5225.
11. Seger, R., Ahn, N. G., Posada, J., Munar, E. S., Jensen, A. M., Cooper, J. A., Cobb, M. H. & Krebs, E. G. (1992) *J. Biol. Chem.* **267**, 14373–14381.
12. Wu, J., Michel, H., Rossomando, A., Haystead, T., Shabanowitz, J., Hunt, D. F. & Sturgill, T. W. (1992) *Biochem. J.* **285**, 701–705.
13. Burnham, C. E., Hawelu-Johnson, C. L., Frank, B. M. & Lynch, K. R. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5605–5609.
14. Pearson, W. & Lipman, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
15. Zeng, D. & Lynch, K. R. (1991) *Mol. Brain Res.* **10**, 219–225.
16. Nakielny, S., Campbell, D. G. & Cohen, P. (1992) *FEBS Lett.* **308**, 183–189.
17. Crews, C. M. & Erikson, R. L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8205–8209.
18. Kozak, M. (1991) *J. Cell Biol.* **115**, 887–903.
19. Kosako, H., Gotoh, Y., Matusda, S., Ishikawa, M. & Nishida, E. (1992) *EMBO J.* **11**, 2903–2908.
20. Hanks, S. & Quinn, A. (1991) *Methods Enzymol.* **200**, 38–62.
21. Nadin-Davis, S. & Nasim, A. (1988) *EMBO J.* **7**, 985–993.
22. Warbrick, E. & Fantes, P. A. (1991) *EMBO J.* **10**, 4291–4299.
23. Boguslawski, G. & Polazzi, J. O. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5848–5852.
24. Teague, M., Chaleff, D. & Errede, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7371–7375.
25. Lindberg, R. A., Quinn, A. M. & Hunter, T. (1992) *Trends Biochem. Sci.* **17**, 114–119.
26. Knighton, D. R., Zheng, J., Eyck, L. F. T., Ashford, V. A., Xuong, N.-h., Taylor, S. S. & Sowadski, J. M. (1991) *Science* **253**, 407–414.
27. Garcia-Bustos, J., Heitman, J. & Hall, M. N. (1991) *Biochim. Biophys. Acta* **1071**, 83–101.
28. Dingwall, C. & Laskey, R. A. (1991) *Trends Biochem. Sci.* **16**, 478–481.