Molecular cloning and sequencing of a carp cDNA encoding mitogen-activated protein kinase kinase

Chang-Jen Huang ^a, Ming-Shyue Lee ^b, Geen-Dong Chang ^b, Fore-Lien Huang ^{a,c}
and Tung-Bin Lo ^{a,b,*}

^a Institute of Biological Chemistry, Academia Sinica, P.O. Box 23-106, Taipei, Taiwan (ROC), ^b Graduate Institute of Biochemical Sciences, Taipei, Taiwan (ROC) and ^c Department of Zoology, National Taiwan University, Taipei, Taiwan (ROC)

(Received 24 September 1993)

Key words: Recombinant DNA; Polymerase chain reaction; MAP kinase kinase; (Carp); (*C. carpio*)

We have cloned a full-length cDNA encoding the mitogen-activated protein kinase kinase (MKK) from a carp liver cDNA library. The cDNA contains 1970 bp with a single open reading frame encoding a 397 amino acid protein. By comparing with known MKK sequences from other species, carp MKK is 78%, 80%, 76% and 58% identical to rat MKK1, rat MKK2, *Xenopus* MKK and *Drosophila* MKK.

Mitogen-activated protein (MAP) kinases, also known as extracellular signal-regulated kinases (ERKs), are a family of serine/threonine kinases that mediate intracellular protein phosphorylation events. They have been shown to participate in a variety of signal transduction pathways and are rapidly activated in response to various extracellular stimuli [1–4]. Upon activation, MAP kinases are phosphorylated on tyrosine and threonine residues [5–8] and then are able to phosphorylate downstream targets on serine and threonine residues. The preferred phosphorylation sites of MAP kinases are identified as Pro-X-Ser/Thr-Pro [9,10]. The downstream substrates of MAP kinases include a wide variety of proteins. Among them are other protein kinases, protein phosphatases, cytoplasmic phospholipase A₂, nuclear transcription factors and proteins engaging in cell cycle activation [11,12]. Therefore, MAP kinase plays an important role in the network of cellular signal transduction pathway.

The activation of MAP kinase is triggered by an upstream activator, called MAP kinase kinase (MKK). MKK is a dual specificity kinase that phosphorylates MAP kinase on both tyrosine and threonine residues [13,14] whereas it is activated and phosphorylated by a further upstream activator on its serine/threonine residues [15]. Recently, MKK has been purified and

cloned from several species, such as fruit fly [16], frog [17], mouse [18], rabbit [19], human [20] and rat [21–23]. All of them are homologous to budding yeast STE7 [24] and to fission yeast byr2 [25]. In order to provide direct evidence for the presence of the MAP kinase kinase in fish and to evaluate the structural evolution of the proteins from lower to higher vertebrates, we cloned and sequenced a full-length cDNA encoding the carp MAP kinase kinase in this study.

Sequences that are identical in *Xenopus* MKK [17], mouse MKK [18], rabbit MKK [19], rat MKK1 [21], *Drosophila* MKK [16] and byr1 [25] were used to design primers for PCR to search for homologous carp MKK. Degenerate oligonucleotides were designed and synthesized from the conserved sequences located in subdomains IV and VII as shown in Fig. 3. With these as primers and cDNA templates derived from polyadenylated RNA from carp liver, a polymerase chain reaction product of 320 base pairs (bp) was obtained (data not shown). The sequence of this product was found to be homologous to those of the known MKKs [16–25]. Therefore this 320 bp cDNA was then used as a probe to screen a carp liver cDNA library, prepared from poly(A)-enriched RNA by unidirectional insertion of cDNA into lambda-ZAP II (purchased from Stratagene). Two positive clones were isolated after screening $5 \cdot 10^5$ plaque-forming units. The pBluescript plasmids containing cDNA inserts were obtained by *in vivo* excision as described by the manufacturer. The resulting plasmids, pMKK4 and pMKK30, contained the cDNA insert of 1.7 and 2.0 kb, respectively. The restriction maps of both plasmids are shown in Fig. 1.

* Corresponding author. Address (a). Fax +886 2 3635038.
The sequence data in this paper have been submitted to the EMBL/Genbank Data Libraries under the accession number L23935.

Fig. 1. The restriction map of pMKK30 and pMKK4. Open box represents coding region and solid line represents 5' (left) and 3' (right) untranslated regions.

The nucleotide sequences of both cDNAs were determined. The cDNA of pMKK4 had the same nucleotide sequence as that of pMKK30, but lacked the first 280 bp in the 5'-flanking region (data not shown). On the basis of the Kozak consensus sequence for initiation codons [26], the starting methionine was predicted to occur at nucleotide 259 of pMKK30 (Fig. 2). With this methionine at the start, the cDNA encodes a protein of 397 amino acids, corresponding to a molecular mass of 45.6 kDa which was similar to those MKKs thus far reported [16–23].

As shown in Fig. 3, carp MKK is homologous to those of other species, with the identity of 78% to rat MKK1 [21], 80% to rat MKK2 [22,23], 76% to *Xenopus* MKK [17], 58% to *Drosophila* MKK [16] and 35% to

fission yeast byr1 [25]. Apparently, the overall structure of MKK is more similar among vertebrates than to that of invertebrates (e.g., *Drosophila*). It is noted that there are five regions of divergence among the aligned MKKs [23]. They are located in the N-terminal region, between kinase domains I and II, between domains V and VI, in a proline-rich looping segment inserted between domains IX and X, and at the C-terminal region. However, the amino acid sequence of carp MKK is more similar to that of rat MKK2 than to those of other MKKs in the divergent region between domains IX and X.

A stretch of N-terminal sequences of all vertebrate MKKs are similar. The N-terminus of rat MKK1 (MPKKKPTPI) is identical to that of *Xenopus* MKK. These sequences have been postulated to resemble the nuclear localization signal of simian virus 40 large tumor antigen [27]. The N-terminal sequences of carp MKK and rat MKK2 are MAPKRRPVPL and MLARRKPVLP, respectively. The functional role of these sequences remains to be investigated.

A potential phosphorylation site at the threonine residue (PSTP) for MAP kinase [9,10] was found at the C-terminal region in all MKKs except yeast MKK (Fig. 3). Such phosphorylation site has been postulated to be involved in the feedback regulation of the signal transduction cascade [19]. This kind of feedback regulation seems to be common for both vertebrate and invertebrate MKKs. Another potential phosphorylation site at the threonine of PRPRTP in the region between domains IX and X was only found in rat MKK1. However, due to point mutation, this phosphorylation site is not found in rat MKK2 and *Xenopus* MKK where the sequences are PRPRPP. In the case of carp MKK, these sequences are PRQRPP. Therefore, this may reflect possible differences in the regulation of different MKKs. In conclusion, we have cloned and sequenced a full-length cDNA encoding carp MKK and this MKK has the very conserved nature, especially to those from the vertebrates.

We thank Dr. C.K. Chou for valuable discussions. This work was supported by grants from the National Science Council (NSC82-0203-B001-119 to C.J.H. and NSC82-0418-B001-15 to T.B.L.) and from the Health Administration (DOH82-HR-C08), Taiwan, Republic of China.

CTCGAGTCGGAGGCCGCTCTCGAAGAGATCCAGCGATGAACGCGCTGAAATAAACCGCAT	60
TTCTCGCAGCTTTGCACTCGCTCTTCCACCTGACACCTAATATTGTTGACAGCGCT	120
CTGGAGGATGACGGGGTGTGAAATAGGGGCTTACGAATCGCAAGAACTTCTATAGAC	180
ACAAGTGTAGCAATTTGAAGGATTGAATGTGAGCAGAGTCAACCTGATGTTCTCTTTGA	240
AATAGATAACAAGCACTATGGCACCAAAAGAAGACAGTCCCTTGATCATAGCCCC	300
MAPKRRPVPLIIAP	14
ACAGGAGAAGGCGAGTCACTAATAGATGACAGCATCAGAGGCCAACCTGGAGGCGCTT	360
TGEGQSTNIDAAASEANLEAL	34
CAGAGGAAGTTAGGGAGCTGGAGTGGATGAACAGCAGAGGAGCGGCTTGAGGCTTTC	420
QRRKLGLDELDEQQRKRLEAF	54
CTCACCAGAAAGCCCAAGTGGGCGAGCTCAAAGATGAAGATTTGACCCCATATGTGAA	480
LTQKKAQVGLKDELDFDPICE	74
CTTGGCGGGCAACGGCGGGGTGGTCCCAAAGTCCGTCACAAAGCCCTCAAGATCGGT	540
L G A G N G V G H K P S R L V	94
ATGGCCAGGAAGCTCATTTCATTTGGAGATCAACACGCGCATCAGGAACAGATCATAGA	600
MARRKLIHLEIKPAIRNQIIR	114
GAGTGCAGGTGCTTCATGAGTGAATTCGCCATACATTTGGGGTTTACGGCGCTTT	660
ELQVLHEECNSPYIVGFYGAFF	134
TACAGCGATGGAGAAATCAGCATCTGTATGGAGCACATGGATGGAGGCTCTTTGGATCAA	720
YSDGEISICMSEHMDGGSLDE	154
GTGCTGAAGAGGCCAGGAGATCCAGAAAGAAATTTGGGCAAGTTAGCATAGCTGTA	780
VLKEARRIEELLGKVLIAV	174
CTCAGAGGTCTGGTATATCTTCGGGAGAAACCAATATATGCACAGACGTTAAGGCC	840
LRGLVYLYLIMHRDIFV	194
TCCAACTCTGGTGAACAGTGGCGGAAATCAAACTGTGGCAGCTTTGGCGTGAGCGG	900
SNILVNSRFLKLCDFGVSG	214
CAGCTCATCGATCCATGGCCAACTCCCTGTTGGAACAGGTCGTACATGTGCGCGGAG	960
QLIDSMAANSFVGTTRSYMTE	234
AGACTCCAGGGCACACTATTCAGTTCACTGGATGTGTGGAGCATGGGCTATCACTG	1020
RLQGTHTYSVQSDVWSMGLSL	254
GTTGAGTGGCTATTGGAGCGCTTCCCATCCCCCTCCCGATGCCAAGGAAGTGGAGGCC	1080
VLEALIGRFPIPPPKALEEA	274
ATATTGGCGGCGAGTGTGGACAAAGTGGGGCAGAAGGCCACGATGTGCCCGAGA	1140
IFGRPVVLDKGGGAEGHSMSPR	294
CAAAGGCTCCAGGACGACAGTGTAGCGGACCGGAATGGACAGCGGCGAGCCATGGCT	1200
QRPPGCRPVSGHGMDSRPAM	314
ATATTGAGCTACTGGATACATTTGCAATGAGCCACCACCAAGCTGCCCATGGCGTC	1260
IFELLDYIVNEPPPKLP HGV	334
TTCACTGATCTTTGAGGAGTTTGATGAAATGCCTCATGAAGAACCGGCTGACAGA	1320
F T T D F E E F V M K C L M K N P A D R	354
GCTGACCTGAAGATGCTAATGGGCCACAGTTTATCAAGCGGCGGAGGTTGAGGAGTG	1380
A D L K M L M G H V F K R A E V	374
GATTTTGCTGGATGGATGTATAAACCATGGGCTTCCAGCGCCAGCACACCACTCAT	1440
D F A G W M C K T M G L P Q P S T H	394
AGTGCTGAATGAACGCGGCTCACAGCGTAGCAACCATGCCATCCTCATCTCTCGCTCT	1500
S A E *	397
ACTTCTCTCTTTCTGCTCTACAGGCGCTAGCGCTGTAAAGCTCCACACCTGTGGG	1560
CTCTTTGGCGTTCTGCTGCTCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1620
TCATAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1680
TTAAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1740
CATCATGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800
CTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1860
ATTTTGGCAGTTAGCATACAGTGTGTTAAATGGATAGCTTGGTTCAAAGGCAACA	1920
TTAATTGAGAATCACTGGCGTTTGATCATCAAGAAAAA	1971

Fig. 2. Nucleotide sequence and deduced amino acid sequence of carp MKK in the single letter code. Numbers of nucleotides and amino acid residues are given on the right.

		MAPKRRPVPLI.IAPFG.EGGST.....NIDAAS.....EANLE	32
r1		M--KK-T-IQ.LN-AP.D-SAV.....-CTSSA...-T--	31
r2		MLARRKPV-L-ALTIM--IA--P-P.....TSEG.....-H-V	35
Xe		M--KK-T-IQ.LN-NP--TAV.....-CTPTA...-T--	36
DSOR1		MS-NKLN--V.LP-VN.TEATVAAATVAPTPPPKTPSGTDLGKPKTSID	47
byr2		MF---RR-N-KGLVLNPNASVK-.....SDNHKEELIN-QK	35
		I* * * * *	
C		ALQRKLGELDLDEQQRKRLEAFLTQKAVQVGLKDEDFDPCIGLAGNGGV	82
r1		--K-K-E-E-----QK-----D-EK-S-----	81
r2		D--K-----K-----D--ER-S-----	85
Xe		--K-K-E-E-----QK-----D-EKVS-----	81
DSOR1		--TET-EG--MGDTE--IKM--S-EKI--S--L-LKVS-----S--	97
byr2		SFESNVAEAPMEQCAHMN-RP-W.....ISD-DNSSLEVVRH--E-----A	79
		* II* * * * * III* * * * * IV* * * * *	
C		VHKVRHKPSRLVMARKLIHLEIKPAIRNQIIRELQVLHCNPSYIVGFG	132
r1		-F-S-----G-----	131
r2		-T-A--R--G-I-----V-----	135
Xe		--S-----TS-I-----	131
DSOR1		-M-----THTH-I-----V-----KK--L--K-----F-H-----	147
byr2		-SL-K-R..NIF-----TVYVGSDSLQK-L--G--H-R-----	129
		* * * * * V* * * * * * * * * *	
C		AFYSIDGEISICMEHMDGGSLDQVLKEARRIPEEILGKVSIAVLRLGLVLR	182
r1		-----K-G-----Q-----IK-T-----	181
r2		-----K-----D-----A-----	185
Xe		-----K-GK--K-----IK-T-----	181
DSOR1		-----Y-----LI--R-G--S-----RTL--K-S-----	197
byr2		--QYKNN--L--Y-C-----AT-R-GGP-LD-----IINSMVK-I-Y	179
		VI* * * * * VII* * * * * * * * * *	
C		EKKHIMHRDVKPSNILVNSRGKEIKLDFGVSGQLIDSMANSFVGTSRYSMS	232
r1		---K-----	231
r2		-----	235
Xe		---K-----	231
DSOR1		DN-A-I-----S-----I-----	247
byr2		NVLH-I--L--VV-----E-VN-V-QT-----ST-----	229
		PERLQ* * * IX* * * * * * * * * *	
C		PARLQGTHTSYVQSDVMSMGLSLVELAIGRPIPPDPAKELEAIFGRPVLD	282
r1		-----I-----M-V-Y-----LL-CQ-EG-----	281
r2		---I-----I-----S-----Y-----	285
Xe		-----I-----M-----Y-----LGCSVGEQDP-----	281
DSOR1		-----I-----M-----MY-----NTAT-S-----A	292
byr2		---IR-GK-T-K-I--L-I-II--TQEL-----	260
		X * * * * *	
C		KGGAEGHSMSPRQRPGRPVSGHGMDSRPAMATIFELLDYIVNEPPPKLPH	332
r1		DAETP.....-P-T-----L-SY-----P-----	327
r2		GADG-P-V-P-----I-----	335
Xe		ASSELA.....-P-----I-SY-P-----P-----	327
DSOR1		DNAE-S.....GQPT-----EPR-----	327
byr2		WSF-NID.....DSIG-LD--HC--Q-E-R---	289
		* * * * * XI * * * * *	
C		GVFTTDFEEFVNMCKMKNPADRADLKMLMGHTFIKRAEVEEVDFAWMCK	382
r1		---SLE-QD-N--I-----E-----Q-V-A--SDA-----L-S-----	377
r2		--SS--Q--N--I-----E--KCLKNPA--S-G-D-----L-R-----	385
Xe		--GAE-QD-N--V-----E-----Q-V-S--QS-L-----L-S-----	377
DSOR1		KI-S-E-KD-DI--K-Q-DE-----T-L-S-PW-RK--L--IS--V--	377
byr2		S.-PE-LRL--DA--H-D-TL--SPQQ-CAMPYFQQ-LMIN--L-S-ASN	338
C		TMGLQPSPSTPHSAE	397
r1		-I--N-----A-SI-----	393
r2		-LR-K-----RT-V-----	400
Xe		--K-----A-GV-----	393
DSOR1		--D--P-----KRNTSPN	392
byr2		FRSS	343

Fig. 3. Alignment of amino acid sequences of carp MKK with those of other species. The reported mouse MKK1 [18], rabbit MKK [19] and human MKK1a [20] are identical to rat MKK1 except for 1, 4 and 3 amino acid substitutions, respectively. Therefore, they are not included for comparison. The positions of PCR primer sites are marked with horizontal arrows. The potential phosphorylation site for MAP kinase is shown by horizontal line above the sequence (PSTP) at the C-terminal region. Abbreviations: dashes, identical amino acids; dots, gaps introduced to maximize alignment; *, amino acids identical in all sequences aligned; I to XI, protein kinase domains; r1, rat MKK1 [21]; r2, rat MKK2 [22,23]; Xe, *Xenopus* MKK [17]; C, carp MKK (this report); DSOR1, *Drosophila* MKK [16]; byr1, protein kinase from *Schizosaccharomyces pombe* [25].

References

- 1 Sturgill, T.W. and Wu, J. (1991) *Biochim. Biophys. Acta* 1092, 350-357.
- 2 Pelech, S.L. and Sanghera, J.S. (1992) *Science* 257, 1355-1356.
- 3 Pelech, S.L. and Sanghera, J.S. (1992) *Trends Biochem. Sci.* 17, 233-238.
- 4 Nishida, E. and Gotoh, Y. (1992) *Int. Rev. Cytol.* 138, 211-238.
- 5 Gomez, N. and Cohen, P. (1991) *Nature* 353, 170-173.
- 6 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.
- 7 Payne, D.M., Rossomando, A.J., Martino, P., Erikson, A.K., Her, J.H., Shabanowitz, J., Hunt, D.F., Weber, M.J. and Sturgill, T.W. (1991) *EMBO J.* 10, 885-892.
- 8 Posada, J. and Cooper, J.A. (1992) *Science* 255, 212-215.
- 9 Clark-Lewis, I., Sanghera, J.S. and Pelech, S.L. (1991) *J. Biol. Chem.* 266, 15180-15184.
- 10 Alvarez, E., Northwood, I.C., Gonzalez, F.A., Birrer, M. and Karin, M. (1991) *J. Biol. Chem.* 266, 15277-15285.
- 11 Ruderman, J.V. (1993) *Curr. Opin. Cell Biol.* 5, 207-213.
- 12 Davis, R.J. (1993) *J. Biol. Chem.* 268, 14553-14556.
- 13 Nakiely, S., Cohen, P., Wu, J. and Sturgill, T.W. (1992) *EMBO J.* 11, 2123-2129.
- 14 Matsuda, S., Kosako, H., Takenaka, K., Moriyama, K., Sakai, H., Akiyama, T., Gotoh, Y. and Nishida, E. (1992) *EMBO J.* 11, 973-982.
- 15 Lange-Carter, C.A., Pleiman, C.M., Gardner, A.M., Blumer, K.J. and Johnson, G.L. (1993) *Science* 260, 315-319.
- 16 Tsuda, L., Inoue, Y.H., Yoo, M.A., Mizuno, M., Hata, M., Lim, Y.M., Adachi-Yamada, T., Ryo, H., Masamune, Y. and Nishida, Y. (1993) *Cell* 72, 407-414.
- 17 Kosako, H., Nishida, E. and Gotoh, Y. (1993) *EMBO J.* 12, 787-794.
- 18 Crews, C.M., Alessandrini, A. and Erikson, R.L. (1992) *Science* 258, 478-480.
- 19 Ashworth, A., Nakiely, S., Cohen, P. and Marshall, C. (1992) *Oncogene* 7, 2555-2556.
- 20 Seger, R., Seger, D., Lozeman, F.J., Ahn, N.G., Graves, L.M., Campbell, J.S., Ericsson, L., Harrylock, M., Jensen, A.M. and Krebs, E.G. (1992) *J. Biol. Chem.* 267, 25628-25631.
- 21 Wu, J., Harrison, J.K., Vincent, L.A., Haystead, C., Haystead, T.A.J., Michel, H., Hunt, D.F., Lynch, K.R. and Sturgill, T.W. (1993) *Proc. Natl. Acad. Sci. USA* 90, 173-177.
- 22 Otsu, M., Terada, Y. and Okayama, H. (1993) *FEBS Lett.* 320, 246-250.
- 23 Wu, J., Harrison, J.K., Dent, P., Lynch, K.R., Weber, M.J. and Sturgill, T.W. (1993) *Mol. Cell. Biol.* 13, 4539-4548.
- 24 Nadin-Davis, S.A. and Nasim, A. (1988) *EMBO J.* 7, 985-993.
- 25 Teague, M.A., Chaleff, D.T. and Errede, B. (1986) *Proc. Natl. Acad. Sci. USA* 83, 7371-7375.
- 26 Kozak, M. (1989) *J. Cell Biol.* 108, 229-241.
- 27 Garcia-Bustos, J., Heitman, J. and Hall, M.N. (1991) *Biochim. Biophys. Acta* 1071, 83-101.