13. T. Espevik, I. S. Figari, G. E. Ranges, M. A. Palladino, Jr., *J. Immunol.* **140**, 2312 (1988).
14. M. Eghbali, *Cell Tissue Res.* **256**, 553 (1989).
15. N. L. Thompson *et al.*, *Growth Factors* **1**, 91 (1988).
16. W. Casscells *et al.*, *Circulation* **80** (suppl. II), 452 (1989).
17. J. R. Gamble and M. A. Vadas, *Science* **242**, 97 (1988).
18. R. F. Furchgott and J. V. Zawadzki, *Nature* **288**, 373 (1980).
19. R. F. Furchgott and P. M. Vanhoutte, *FASEB J.* **3**, 2007 (1989).
20. N. Aoki, H. Bitterman, M. E. Brezinski, A. M. Lefer, *Br. J. Pharmacol.* **95**, 735 (1988).
21. R. M. J. Palmer, A. G. Ferrige, S. Moncada, *Nature* **327**, 524 (1987).
22. L. J. Ignarro, *Circ. Res.* **65**, 1 (1989).
23. G. M. Rubanyi and P. M. Vanhoutte, *Am. J. Physiol.* **250**, H822 (1986).
24. J. M. Downey, D. J. Hearse, D. M. Yellon, *J. Mol. Cell. Cardiol.* **20** (suppl 2), 55 (1988).
25. J. M. McCord and I. Fridovich, *J. Biol. Chem.* **244**, 6049 (1969).
26. K. J. Tracey *et al.*, *Science* **234**, 470 (1986).
27. N. Aoki, M. Siegfried, A. M. Lefer, *Am. J. Physiol.* **256**, H1509 (1989).
28. J. K. Kjekshus and B. E. Sobel, *Circ. Res.* **27**, 403 (1970).
29. D. Chantry, M. Turner, E. Abney, M. Feldmann, *J. Immunol.* **142**, 4295 (1989).
30. S. Tsunawaki, M. Sporn, A. Ding, C. Nathan, *Nature* **334**, 260 (1988).
31. R. Derynck *et al.*, *ibid.* **316**, 701 (1985).
32. C. Lucas, B. M. Fendly, V. R. Mukku, W. L. Wong, M. A. Palladino, in *Methods in Enzymology*, part C, *Peptide Growth Factors*, D. Barnes, J. Mather, G. H. Sato, Eds. (Academic Press, New York, in press).
33. D. Pennica, J. S. Hayflick, T. Bringman, M. A. Palladino, D. V. Goeddel, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6060 (1985).
34. S. M. Kramer and M. E. Carver, *J. Immunol. Methods* **93**, 201 (1986).
35. We thank P. Taylor III for technical assistance during these investigations.

13 February 1990; accepted 19 April 1990

An Insulin-Stimulated Protein Kinase Similar to Yeast Kinases Involved in Cell Cycle Control

TERI G. BOULTON, GEORGE D. YANCOPOULOS, JILL S. GREGORY, CLIVE SLAUGHTER, CAROLYN MOOMAW, JOAN HSU, MELANIE H. COBB*

A protein kinase characterized by its ability to phosphorylate microtubule-associated protein-2 (MAP2), is thought to be an early intermediate in an insulin-stimulated phosphorylation cascade and in a variety of other mammalian cell responses to extracellular signals. A complementary DNA that encodes this protein serine-threonine kinase has been cloned, and the protein designated extracellular signal-regulated kinase 1 (ERK1). ERK1 has striking similarity to two protein kinases, KSS1 and FUS3, from yeast. The yeast kinases function in an antagonistic manner to regulate the cell cycle in response to mating factors. Thus, ERK1 and the two yeast kinases constitute a family of evolutionarily conserved enzymes involved in regulating the response of eukaryotic cells to extracellular signals.

INSULIN AND SEVERAL OTHER EXTRACELLULAR cues initiate signal transduction by activating a tyrosine kinase either intrinsic to or associated with their specific receptors. These activated kinases then trigger phosphorylation cascades consisting of protein serine-threonine kinases, few of which have been identified. MAP2 kinase is rapidly stimulated in response to insulin (*1–4*), and is one of the few serine-threonine protein kinases known to be phosphorylated on tyrosine residues in vivo (*2*). It can phosphorylate an insulin-sensitive ribosomal protein S6 kinase in vitro causing an increase in its activity (*5, 6*). Thus, the MAP2 kinase may be an early intermediate regulating insulin-dependent protein kinase cascades and may be a direct substrate of the tyrosine kinase of the insulin receptor. The finding that phosphorylation on threonine as well as tyrosine residues is required for MAP2 kinase activity (*7, 8*) suggests that it is regulated by multiple phosphorylations. Other kinases poised at critical steps in phosphorylation cascades, such as Raf (*9*) and cdc2$^+$ (*10*), also display this behavior.

In addition to stimulation by insulin, MAP2 kinase activity can be rapidly increased by a variety of other extracellular signals that promote cellular proliferation and/or differentiation (*1–4*). In this regard, MAP2 kinase may be equivalent to pp42 (*11*), a protein whose phosphotyrosine content increases after exposure to growth factors and transformation by viruses (*12*). MAP2 kinase activity is also stimulated in post-mitotic adrenal chromaffin cells in response to signals that induce catecholamine secretion (*13*), and in PC12 cells in response to nerve growth factor–induced neuronal differentiation (*14*). Thus, MAP2 kinase is likely to be involved in many different signal transduction pathways in a wide variety of cell types.

Purified MAP2 kinase consists of one major band with a molecular size of 43,000 daltons (*8*). After tryptic cleavage of this band, amino acid sequences of seven peptides were obtained (*15*). Although none of the peptide sequences are in the GenBank database, consensus sequences characteristic of serine-threonine protein kinases (*16*), GEGAYG (part of the nucleotide binding site) and DLKPSN, were found (*17*). A series of degenerate oligonucleotides corresponding to different regions of the amino acid sequence were used for polymerase chain reactions (PCR) (*18*). Complementary DNA templates prepared from Rat 1 fibroblasts and oligonucleotides corresponding to segments of the peptides that contain the conserved GEGAYG and DLKPSN sequences were used in a PCR and yielded an amplified DNA fragment of an expected size. Based on similarities to other protein kinases, we assumed that of the two sequences GEGAYG is closer to the NH_2-terminus of the protein; the two sequences are separated by about 120 amino acids in most protein kinases. PCR with the degenerate oligonucleotide encoding the amino acids QYIGEG (5′-TTCTAGAATTC CA-(A,G) TA(C,T) AT(A,T,C) GG(A,T,C,G) GA(A,G) GG-3′) and the degenerate oligonucleotide corresponding to the noncoding strand for the amino acids DLKPSN (5′-TTCTCGAGTCGAC (A,G)TT (A,T,C,G)-GA (A,T,C,G)GG (C,T)TT (A,T,C,G)-A(A,G) (A,G)TC-3′) yielded an amplified product of ~360 base pairs (bp). The PCR product was sequenced to confirm that it encoded part of a novel protein kinase, and used as a probe to screen a rat brain cDNA library (*19*). A single clone, which hybridized at high stringency, was isolated and contained a 1.9-kb cDNA insert.

A probe made from the cDNA insert identified a 1.9-kb transcript by Northern (RNA) blot analysis, indicating that our cDNA clone contains an insert that is a nearly full-length transcript. The ERK1 transcript was detectable in all tissues and cell lines examined (Fig. 2). The highest concentrations (three to six times higher) are in the central nervous system (brain and

T. G. Boulton, J. S. Gregory, M. H. Cobb, Department of Pharmacology, University of Texas Southwestern Graduate School of Biomedical Sciences, Dallas, TX 75235.
G. D. Yancopoulos, Regeneron Pharmaceuticals, Inc., Tarrytown, NY 10591.
C. Slaughter, C. Moomaw, J. Hsu, Department of Biochemistry and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75235.

*To whom correspondence should be addressed.

spinal cord). The transcript was detectable in Rat 1 fibroblasts and PC12 cell lines, both of which express acutely regulated MAP2 kinase activity.

The cDNA insert has a single long open reading frame that encodes at least a protein of 367 residues with a molecular size of 42,038 daltons. It contains the primary sequence of seven of the tryptic peptides isolated from the insulin-stimulated MAP2 kinase. In addition, the exact sequences of two additional peptides from a mixed sequence were found in the translated cDNA. Together these tryptic peptides (Fig. 1) consisted of 115 residues all of which are in accord with the translated cDNA, accounting for more than 27% of the putative translation product. The exact correspondence between the extensive tryptic peptide sequence and the predicted translation product provides substantial evidence that the cDNA described herein encodes the insulin-stimulated MAP2 kinase. Because of its mode of regulation, we have named the protein encoded by this cDNA extracellular signal-regulated kinase 1 (ERK1).

ERK1 contains the 15 nearly invariant residues noted in protein kinases as well as similarity with all of the defined subdomains (*16*). In addition, ERK1 contains residues characteristic of serine-threonine protein kinases. A comparison with other protein kinases reveals striking similarities between ERK1 and two yeast kinases (Fig. 3), KSS1 (*20*) and FUS3 (*21*), which mediate the yeast response to pheromones (see below). ERK1 is 51% identical to KSS1 and 52% identical to FUS3 over their shared lengths. KSS1 and FUS3 display about the same degree of similarity (54% identical) to each other as they do to ERK1; they are all less closely related to other kinases than they are to each other. The significance of the similarity is suggested by previous observations that mammalian kinases may contain no more than 50 to 65% identity with their yeast homologs (*22*). ERK1, KSS1, and FUS3 share their next most impressive similarities with the CDC28/cdc2$^+$ subfamily of kinases. ERK1 is ~38% identical to human CDC2$^+$ kinase over their shared lengths; KSS1 is ~40% identical and FUS3 is ~35% identical to the yeast CDC28/cdc2$^+$ kinases. However, the three lack the VPSTAIR sequence found in subdomain III of all CDC28 functional homologs. Furthermore, the three also share COOH-terminal extensions not found on CDC28/cdc2$^+$. Unlike FUS3 and KSS1, ERK1 contains a significant NH_2-terminal extension of at least 67 amino acids.

ERK1 also differs from its yeast homologs at the COOH-terminus and between the DFG and APE motifs of subdomains VII and VIII, which contain inserts of different lengths with phosphorylatable residues (Thr-186 and Thr-190 in ERK1). Both regions, poorly conserved between the yeast kinases, have been implicated in determining functional characteristics of individual kinases. In a number of kinases, the segment between subdomains VII and VIII

```
                                  Glu Pro Arg Gly Thr Ala Gly Val Val Pro Val Val Pro Gly Glu Val Glu Val Val
1                                 GAG CCC AGG GGA ACT GCT GGG GTC GTC CCG GTG GTC CCC GGG GAG GTG GAG GTG GTG
                                              *                 *                 *

   Lys Gly Gln Pro Phe Asp Val Gly Pro Arg Tyr Thr Gln Leu Gln Tyr Ile Gly Glu Gly Ala Tyr Gly Met Val Ser Ser
58 AAG GGG CAG CCA TTC GAC GTG GGC CCA CGC TAC ACG CAG CTG CAG TAC ATC GGC GAG GGC GCG TAC GGC ATG GTC AGC TCA
                                        *   *

   Ala Tyr Asp His Val Arg Lys Thr Arg Val Ala Ile Lys Lys Ile Ser Pro Phe Glu His Gln Thr Tyr Cys Gln Arg Thr
139 GCA TAT GAC CAC GTG CGC AAG ACC AGA GTG GCT ATC AAG AAG ATC AGC CCC TTC GAG CAT CAA ACC TAC TGT CAG CGC ACG
          *

   Leu Arg Glu Ile Gln Ile Leu Leu Gly Phe Arg His Glu Asn Val Ile Gly Ile Arg Asp Ile Leu Arg Ala Pro Thr Leu
220 CTG AGA GAA ATC CAG ATC TTG CTC GGA TTC CGC CAT GAG AAT GTC ATA GGC ATC CGA GAC ATC CTC AGA GCA CCC ACC CTG

   Glu Ala Met Arg Asp Val Tyr Ile Val Gln Asp Leu Met Glu Thr Asp Leu Tyr Lys Leu Leu Lys Ser Gln Gln Leu Ser
301 GAA GCC ATG AGA GAT GTT TAC ATT GTT CAG GAC CTC ATG GAG ACG GAC CTG TAC AAG CTG CTA AAG AGC CAG CAG CTG AGC
                                                                                                        *

   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
382 AAT GAC CAC ATC TGC TAC TTC CTC TAC CAG ATC CTC CGG GGC CTC AAG TAC ATA CAC TCG GCC AAT GTG CTG CAC CGG GAC
                  *                                                 *   *                 *

   Leu Lys Pro Ser Asn Leu Leu Ile Asn Thr Thr Cys Asp Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Ile Ala Asp Pro
463 CTG AAG CCC TCC AAT CTG CTT ATC AAC ACC ACC TGC GAC CTT AAG ATC TGT GAT TTT GGC CTT GCC CGG ATT GCT GAC CCT
                                                                  *   *   *

   Glu His Asp His Thr Gly Phe Leu Thr Glu Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Ile Met Leu Asn Ser Lys
544 GAG CAC GAC CAC ACT GGC TTT CTG ACC GAG TAT GTG GCC ACA CGC TGG TAC CGA GCC CCA GAG ATC ATG CTT AAC TCC AAG
                        *                 *

   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
625 GGC TAC ACC AAA TCC ATT GAC ATC TGG TCT GTG GGC TGC ATT CTG GCT GAG ATG CTC TCC AAC CGG CCT ATC TTC CCC GGC

   Lys His Tyr Leu Asp Gln Leu Asn His Ile Leu Gly Ile Leu Gly Ser Pro Ser Gln Glu Asp Leu Asn Cys Ile Ile Asn
706 AAG CAC TAC CTG GAC CAG CTC AAC CAC ATT CTA GGT ATA CTG GGT TCC CCA TCC CAA GAG GAC CTA AAT TGT ATC ATT AAC

   Met Lys Ala Arg Asn Tyr Leu Gln Ser Leu Pro Ser Lys Thr Lys Val Ala Trp Ala Lys Leu Phe Pro Lys Ser Asp Ser
787 ATG AAG GCC CGA AAC TAC CTA CAG TCT CTG CCC TCT AAA ACC AAG GTG GCT TGG GCC AAG CTT TTT CCC AAA TCT GAC TCC
                                                    *

   Lys Ala Leu Asp Leu Leu Asp Arg Met Leu Thr Phe Asn Pro Asn Lys Arg Ile Thr Val Glu Glu Ala Leu Ala His Pro
868 AAA GCT CTT GAC CTG CTG GAC CGG ATG TTA ACC TTT AAC CCA AAC AAG CGC ATC ACA GTA GAG GAA GCA CTG GCT CAC CCT

   Tyr Leu Glu Gln Tyr Tyr Asp Pro Thr Asp Glu Pro Val Ala Glu Glu Pro Phe Thr Phe Asp Met Glu Leu Asp Asp Leu
949 TAC CTG GAA CAG TAC TAT GAT CCG ACA GAT GAA CCA GTG GCT GAG GAG CCA TTC ACC TTT GAC ATG GAG CTG GAT GAT CTC

   Pro Lys Glu Arg Leu Lys Glu Leu Ile Phe Gln Glu Thr Ala Arg Phe Gln Pro Gly Ala Pro Glu Ala Pro END CAAGAAC
1030 CCC AAG GAG CGG CTG AAG GAG CTG ATC TTC CAA GAG ACA GCC CGC TTC CAG CCA GGG GCA CCA GAG GCC CCC TAA

1112 AGACACCCCTGTCCTTTTGGACCTGGTCTGCTCTACCTGCTCCTTCTCTGCAGATTGTTAGAAAATGAACTTTGCTCAACCCGGACCCCGGCAGCCCAGGCTGGACC

1219 AAGGGTGGGCCTGGCACCCCTCTCACTCTGCTGGGGTCTCCTCGTTCAAGAGGCTTCTCCCACTCCAGTCCCCTGCCCCATCTCCCCTTGACCTGAGTGATGAGGTG

1326 GTCCCAGAGCTGATCTCTGCTGCTGTGTCTTTATCTATCCCTGCTAGCCCCAGCTCTGGTAGACGGTTCTGGAATGGAAGGGCTATGACCGCCCTAGGACCTGTGCT

1433 ACAGAGGGGTGGAGGGCACTGAGTAGGCTAAGCTCTGCCCTACTCATCCTGTTGGAACCCCACCCCATTTTCCCTGACAGAACATTCCTAAATCTCAAGGGCTAGTT

1540 TCCCTGAGGAGCCAGCCTAGGCCTAACCCTCTCCCTCTCAAGCTGCCACATGTAACGCCCTTGCTGCTTCTGTGTGTGGGTGATTGGATGTGGAGGCGGGGCCCGTG

1647 GAGAGCCCGTGCCCCTCCCCACCTCCCTGTGCCTGTATCTAATATATAAATATAGAGATGTGTATATGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Fig. 1. The complete nucleotide sequence of the ERK1 cDNA and its predicted protein product. Nucleotide sequencing was performed as described (*28*). Asterisks denote the residues most conserved among all protein kinases. The open reading frame extends to the end of the nearly full-length cDNA insert; the size of the purified kinase predicts that the initiator methionine of ERK1 is located just upstream of the sequence shown here. The nine tryptic peptides that were sequenced are underlined. All residues determined by amino acid sequencing matched the cDNA-encoded protein sequence; questionable residues were verified from the cDNA encoded protein sequence. The fourth and seventh peptides indicated represented the minor peptide components described in the text.

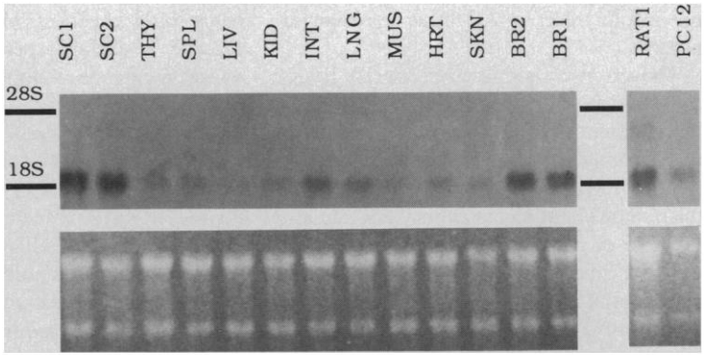

Fig. 2. Expression of ERK1 transcript in rat tissues and cell lines. The QYDL fragment (*19*), directly amplified from the ERK1 cDNA clone was used as the probe in the analysis depicted; several other probes prepared from the ERK1 cDNA yielded identical expression patterns. RNA isolation, Northern blotting, and ^{32}P-radiolabeling by PCR and hybridization to labeled probes were performed as described (*27*). The ethidium bromide–stained gel is shown to allow comparison of the total amount of RNA per sample. The abundance of ERK1 transcripts in the adult rat brain was estimated to be ~0.0005% based on screening 10^6 phage from a rat brain cDNA library at high stringency. SC1, postnatal day 19 spinal cord; SC2, adult spinal cord; thy, thymus; spl, spleen; liv, liver; kid, kidney; int, intestine; lng, lung; mus, muscle; hrt, heart; skn, skin; BR2, adult brain; BR1, postnatal day 10 brain; Rat1, Rat 1 fibroblasts; PC12, pheochromocytoma cells.

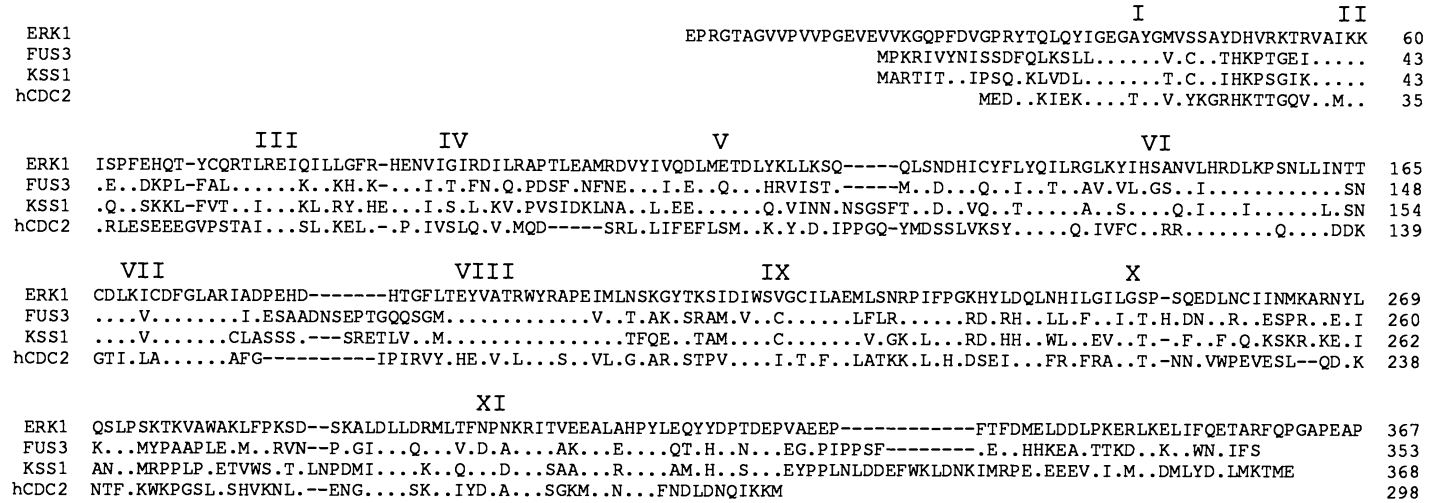

Fig. 3. Alignments of ERK1, FUS3, KSS1, and human cdc2$^+$ protein sequences. Roman numerals designate subdomains conserved in protein kinases. Dots indicate identities to the ERK1 sequence, dashes indicate spaces introduced to maximally align sequences. Alignments were computer-generated (MacVector Computer Analysis Software, International Biotechnologies, Inc., New Haven, Connecticut) and then visually maximized. Percent identities provided in the text were determined over the length of the cdc2$^+$ molecule; residues in insertions were ignored for these calculations.

is autophosphorylated in a manner that influences enzymatic activity [cAMP-dependent protein kinase (*23*), insulin receptor (*24*)]. Conservation among the kinases may also reveal functionally important residues. There is a conserved tyrosine in subdomain I, whose phosphorylation in CDC28 inhibits its protein kinase activity (*25*). A sequence similar to the regulatory autophosphorylated region of the insulin receptor exists near the COOH-terminus of ERK1 (residues 315 to 327) and may be a potential site of tyrosine phosphorylation. Three tyrosine residues located within this region are spaced identically to those in the insulin receptor (YX_3YY); four acidic residues, important determinants for recognition by protein tyrosine kinases, are located nearby. The presence of sequences resembling the insulin receptor phosphorylation site is consistent with other evidence suggesting that ERK1 (MAP2 kinase) is a substrate for the insulin receptor or other tyrosine kinases.

The extensive similarity between ERK1 and the two yeast kinases, KSS1 and FUS3, is consistent with an evolutionarily conserved function for this new family of kinases in mediating the response to extracellular signals and suggests that ERK1, as a member of this family, may promote entry into the cell cycle. The yeast kinases play antagonistic roles in regulating the yeast cell cycle in response to mating factors, the only known peptide hormones that mediate intercellular communication in yeast. Both kinases seem to act by fine-tuning the activity of CDC28, a related protein serine-threonine kinase, which is the indispensable regulator of the mitotic cycle, probably via interactions with a yeast cyclin. FUS3 apparently has two regulatory functions (*21*). It seems to be important in leading to pheromone-induced cell cycle arrest in G1, either by directly inhibiting the activation of CDC28 or by promoting the inactivation of a cyclin required for activation of CDC28. Activation of FUS3 by pheromones also independently promotes mating-specific functions. By contrast, KSS1 promotes reentry into the cell cycle after pheromone-induced cell cycle arrest (*20*); KSS1 may function by activating the same cyclin that FUS3 may inactivate. ERK1 apparently represents a mammalian counterpart, perhaps functionally as well as evolutionarily, to the yeast kinases. Thus, ERK1 may act via similar pathways to regulate the cell cycle in response to extracellular signals. As with FUS3, ERK1 may also function in a regulatory capacity in postmitotic cells and in responses that do not directly involve the cell cycle. The extensive similarity between ERK1 and the yeast kinases raises the possibility that yeast may provide a useful experimental system in which to introduce cloned ERK1 for functional analysis.

We have also cloned other kinases highly related to ERK1 (*26*), one of which is even more closely related to KSS1 and FUS3 than is ERK1. The identification of a mammalian family of ERK-related enzymes, which are structural homologs of KSS1 and FUS3, suggests that multicellular, higher eukaryotes have appropriated kinases, originally utilized in a primordial mating factor response by unicellular organisms, to mediate responses to extracellular signals. The molecular cloning of this family will facilitate the elucidation of their mechanisms of regulation and allow analysis of their physiological functions.

REFERENCES AND NOTES

1. L. B. Ray and T. W. Sturgill, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1502 (1987).
2. ———, *ibid.* **85**, 3753 (1988); *J. Biol. Chem.* **263**, 12721 (1988).
3. M. Hoshi, E. Nishida, H. Sakai, *J. Biol. Chem.* **263**, 5396 (1988).
4. T. G. Boulton, J. S. Gregory, S.-M. Jong, L.-H. Wang, L. Ellis, M. H. Cobb, *ibid.* **265**, 2713 (1990).
5. T. W. Sturgill, L. B. Ray, E. Erikson, J. L. Maller, *Nature* **334**, 715 (1988).
6. J. S. Gregory, T. G. Boulton, B.-C. Sang, M. H. Cobb, *J. Biol. Chem.* **264**, 18397 (1989).
7. N. G. Anderson, J. L. Maller, N. K. Tonks, T. W. Sturgill, *Nature* **343**, 651 (1990).
8. T. G. Boulton, J. S. Gregory, M. H. Cobb, unpublished data.
9. D. Morrison, D. Kaplan, V. Rapp, T. Roberts, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8855 (1988).
10. G. Draetta and D. Beach, *Cell* **54**, 17 (1988).
11. J. A. Cooper and T. Hunter, *Mol. Cell. Biol.* **1**, 165 (1981).
12. A. Rossomando, M. Payne, M. Weber, T. Sturgill, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6940 (1989).
13. C. Ely, K. Oddie, J. Litz, A. Rossomando, S. Kanner, T. Sturgill, S. Parsons, *J. Cell Biol.* **110**, 731 (1990).
14. H. Tsao, J. Aletta, L. Greene, *Soc. Neurosci. Abstr.* **15**, 869 (1989); T. Miyasaka, M. Chao, P. Sherline, A. Saltiel, *J. Biol. Chem.* **265**, 4730 (1990).
15. MAP2 kinase was purified from Rat 1 HIRc B cells (obtained from D. McClain, Veterans Administration Medical Center, San Diego, CA), treated with insulin (from M. Root, Eli Lilly). The major component (43,000 daltons) was further purified by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) and digested with trypsin. The resulting peptides were subjected to high-performance liquid chromatography (*8*) and one of the resulting peaks was subjected to a second chromatographic separation. Amino acid sequence was obtained from seven distinct peaks. One peak contained a mixture of three peptides, with one major and two minor components; the sequence of the major peptide was determined on the basis of recovery, but the assignment of the amino acids in the minor components to their respective peptide sequences and the verification of a few questionable residues (Cys) were based on the cDNA sequence (see legend to Fig. 1).
16. S. K. Hanks, A. M. Quinn, T. Hunter, *Science* **241**, 42 (1988).
17. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro;

Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

18. A series of degenerate oligonucleotides were synthesized that correspond to the coding or noncoding strands for segments of the tryptic peptide sequences. The oligonucleotides contained nondegenerate 5′ ends; the end of each coding strand oligonucleotide had an Eco RI restriction site, while the end of each noncoding strand oligonucleotide contained a Sal I restriction site. Each coding strand oligonucleotide was combined with each noncoding oligonucleotide in individual PCR reactions with cDNA from Rat 1 cells as template; the PCR reactions and the preparation of the genomic and cDNA templates were performed as described (*27*) [A. Bothwell, G. Yancopoulos, F. Alt, *Methods for Cloning and Analysis of Eukaryotic Genes* (Jones and Bartlett, Boston, MA, 1990)]. The amplified product obtained with the QYIGEG coding oligonucleotide and the DLKPSN noncoding oligonucleotide (designated QYDL) was isolated on a Sephadex G-50 spin column, digested with Eco RI and Sal I, gel-purified with 2% Nusieve (FMC Bioproducts), and subcloned into the pGEM4Z vector (Promega).
19. Plaques (600,000) from a rat brain cDNA library constructed in the Lambda Zap II vector (Stratagene) were screened with the subcloned QYDL PCR product as the probe, radiolabeled with a PCR-based protocol (*27*). Hybridization conditions have been described (*27*); after hybridization the library filters were washed first at low stringency [2× SSC (20 mM sodium citrate, pH 7.0, 0.15 M NaCl), 0.1% SDS, 60°C] and then at high stringency (0.2× SSC, 0.1% SDS, 60°C).
20. W. E. Courchesne, R. Kunisawa, J. Thorner, *Cell* **58**, 1107 (1989).
21. E. A. Elion, P. L. Grisafi, G. R. Fink, *ibid.* **60**, 649 (1990).
22. M. G. Lee and P. Nurse, *Nature* **327**, 31 (1987).
23. L. R. Levin and M. J. Zoller, *Mol. Cell. Biol.* **10**, 1066 (1990).
24. L. Ellis *et al.*, *Cell* **45**, 721 (1986).
25. K. Gould and P. Nurse, *Nature* **342**, 39 (1989).
26. T. Boulton, M. H. Cobb, G. Yancopoulos, unpublished data.
27. C. Maisonpierre *et al.*, *Science* **247**, 1446 (1990).
28. Sequencing was performed with the dideoxynucleotide chain termination method [F. Sanger, S. Nicklen, A. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977)], with the Sequenase Kit (version 2.0), and recommended protocols (U.S. Biochemicals). All sequences were verified by sequencing both strands of the DNA with appropriate oligonucleotides corresponding to the ERK1 sequence or flanking plasmid sequence.
29. We thank L. S. Schleifer, and the rest of the Regeneron community (especially M. Furth and S. Squinto) for their support and assistance, A. Gilman for discussions, and J. Hicks for preparation of the manuscript. Supported by Regeneron Pharmaceuticals, Inc., and grant DK 34128, Research Career Development Award DK 01918 from the NIH, and American Cancer Society institutional grant IN-142 to M.H.C. T.G.B.'s work is in partial fulfillment of the requirements for the Ph.D. degree.

17 May 1990; accepted 31 May 1990

Recognition of a Peptide Antigen by Heat Shock–Reactive γδ T Lymphocytes

WILLI BORN, LISA HALL, ANGELA DALLAS, JOEL BOYMEL, THOMAS SHINNICK, DOUGLAS YOUNG, PATRICK BRENNAN, REBECCA O'BRIEN

Small synthetic peptides that correspond to different portions of the 65-kilodalton mycobacterial heat shock protein (Hsp65) were used to identify a putative antigenic epitope for γδ cells. Weaker γδ responses to the equivalent portion of the autologous homolog, mouse Hsp63, were also seen. The stimulatory epitope overlaps with an epitope recognized by arthritogenic αβ T cell clones. The data suggest that γδ cells have a role in autoimmune disorders and imply that these cells recognize ligands by a mechanism similar to that of αβ T lymphocytes, that is, in the form of small processed protein fragments bound to antigen-presenting molecules.

LYMPHOCYTES BEARING THE αβ T cell receptor (TCR) recognize foreign antigens either in the form of small, processed protein fragments or as "superantigens" (*1*). The nature of the ligands recognized by γδ cells is not known, although structural similarities between αβ and γδ TCRs, and the demonstration of γδ cell responses to major histocompatibility complex (MHC) class I and class I like alloantigens, including Qa, TL, and CD1, suggest that γδ and αβ ligands are similar (*2, 3*). The smaller receptor repertoire of γδ cells implies, however, that fewer ligands are recognized. For these reasons and because of the frequent association of γδ cells with epithelia, it has been suggested that γδ cells may act as a first line of defense (*4*) and respond to stress signals of surrounding tissues (*5*) rather than recognizing diverse foreign antigens directly.

Some γδ cells respond to mycobacterial antigens (*6–9*). Many of these can be stimulated with purified protein derivative (PPD) of *Mycobacterium tuberculosis*, containing the presumably denatured mycobacterial 65-kD heat shock protein (Hsp65), and with purified recombinant Hsp65 (*6–8*). Mycobacterial Hsp65 is known to be an immunodominant antigen (*10*) that contains many stimulatory epitopes for B cells and αβ T cells. The αβ T cell epitopes have been defined with synthetic peptides correspond-

W. Born, Department of Microbiology and Immunology, University of Colorado Health Sciences Center, and Department of Medicine, National Jewish Center for Immunology and Respiratory Medicine, Denver, CO 80206.
L. Hall, A. Dallas, J. Boymel, R. O'Brien, Department of Medicine, National Jewish Center for Immunology and Respiratory Medicine, Denver, CO 80206.
T. Shinnick, Division of Bacterial Diseases, Centers for Disease Control, Atlanta, GA 30333.
D. Young, Medical Research Council Tuberculosis and Related Infections Unit, Hammersmith Hospital, Ducane Rd, London W12 OHS, United Kingdom.
P. Brennan, Department of Microbiology, Colorado State University, Fort Collins, CO 80523.

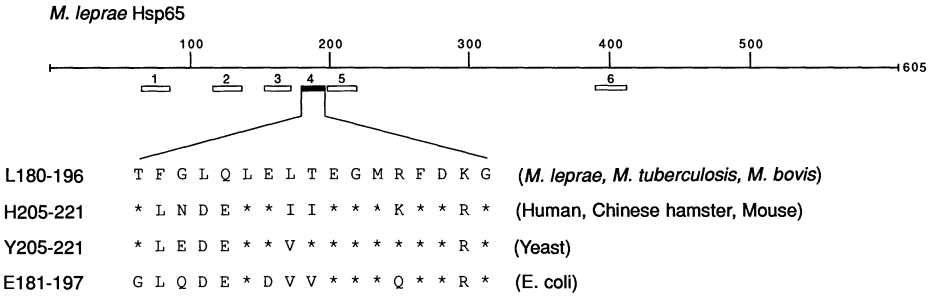

Fig. 1. Linear map of *M. leprae* Hsp65 (*13*) and synthetic heat shock peptides. The six synthetic peptides shown (*12*) and nine further peptides listed in Fig. 2 (*11, 12*) were used to map a putative stimulatory epitope for γδ cells. Also shown are the sequences of newly synthesized peptides, corresponding to the same epitope in homologous proteins of other species (*13, 19*). Peptides are designated by a capital letter to indicate the species origin of the protein sequence (L for *M. leprae*, *M. tuberculosis*, and *M. bovis*; H for human, Chinese hamster, and mouse; Y for yeast; and E for *E. coli*) and by the position numbers of the NH_2-terminal and COOH-terminal amino acids (*13*). New peptides were synthesized according to the solid-phase methods generally described by Merrifield *et al.* (*22*) and detailed by Buus *et al.* (*23*). To verify primary structures, we purified peptides by high-performance liquid chromatography (HPLC) and analyzed for amino acid compositions and sequences (Applied Biosystems model 470A gas-phase sequencer). The predicted protein sequence of a Hsp65 homolog in mice was derived by polymerase chain reaction amplification and sequencing of DNA complementary to mouse mRNA, with the use of primers corresponding to short stretches of gene sequence identical in Chinese hamsters and humans (*19*).