Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 6035–6039, September 1987
Biochemistry

Two yeast genes that encode unusual protein kinases

(oligonucleotide probes/conserved sequences/*KIN* genes/tyrosine kinases)

DAVID E. LEVIN, CHARLOTTE I. HAMMOND*, ROBERT O. RALSTON†, AND J. MICHAEL BISHOP

Department of Microbiology and Immunology and The G. W. Hooper Research Foundation, University of California, San Francisco, CA 94143

Contributed by J. Michael Bishop, April 27, 1987

ABSTRACT **Mixed synthetic oligonucleotides encoding sequences conserved among tyrosine-specific protein kinases were used to probe the genome of the budding yeast *Saccharomyces cerevisiae*. Two genes with homology to protein kinases were isolated and characterized by DNA sequence analysis. These genes, designated *KIN1* and *KIN2*, are closely related to each other. Among previously characterized protein kinases, the products of *KIN1* and *KIN2* are most closely related to the bovine cAMP-dependent protein kinase (30% amino acid identities) and the protein encoded by the v-*src* oncogene (27% and 25% identities with *KIN1* and *KIN2*, respectively) within their putative kinase domains. *KIN1* and *KIN2* are transcribed into 3.5-kilobase mRNAs that contain uninterrupted open reading frames encoding polypeptides of 117 kDa and 126 kDa, respectively. The predicted proteins are unusual in two respects: (*i*) their catalytic domains are carried near the N termini of relatively large proteins, in contrast to the majority of characterized protein kinases, and (*ii*) these catalytic domains are structural mosaics, with some features characteristic of tyrosine-specific protein kinases and other elements that are distinctive of serine/threonine-specific enzymes.**

Protein kinases comprise a diverse group of regulatory proteins widely suspected to be involved in growth control and malignant transformation in eukaryotes (for reviews, see refs. 1–3). These enzymes are thought to exert their regulatory effects by phosphorylating other proteins. Identification of important substrates for the kinases is the key to understanding the physiological significance of protein phosphorylation. To study the role of protein kinases in normal and malignant cells, it will be useful to take advantage of the existence of these enzymes in species that lend themselves to rigorous genetic analysis. The budding yeast *Saccharomyces cerevisiae* is supremely amenable to such investigation.

The number of unique protein kinases that have been described now exceeds 50. These can be classified into two subfamilies based on substrate specificity. Some specifically phosphorylate tyrosine residues, whereas others modify serine and threonine residues (3). Members of these subfamilies can be distinguished by characteristic structural features that are highly conserved within each subfamily. Several protein kinase genes have been isolated from yeast (4–8), but all have been shown to be, or are, predicted from their amino acid sequences to be serine/threonine-specific enzymes. Although a tyrosine-specific protein kinase gene has yet to be found in the yeast genome, recent biochemical evidence suggests that a low level of tyrosine kinase activity is detectable in crude extracts of yeast cells (9). We have approached the isolation of tyrosine-specific protein kinase genes from yeast by searching for sequences that are highly conserved within this subfamily.

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.

MATERIALS AND METHODS

Genomic Yeast Library. The bacteriophage λ library of genomic yeast DNA, provided by M. Snyder (Stanford University), was constructed by cloning a partial *Sau*3A digest of yeast DNA into vector EMBL3a. Since this vector accommodates inserts of size 10–15 kilobases (kb), and the genome size of yeast is 10^7 base pairs (bp) (10), 1000 plaques correspond to approximately one genome equivalent of yeast DNA.

Hybridization Analyses. Genomic yeast DNA was prepared as described by Winston *et al.* (11). DNA was digested with restriction endonucleases; then fragments were separated by agarose gel electrophoresis and transferred to nitrocellulose (12). Yeast poly(A)$^+$ RNA was isolated from a cells and prepared for hybridization as described (13). The phage library was prepared for plaque hybridization on nitrocellulose filters (14). The oligonucleotide pools were synthesized by the methoxy phosphoramidite method (15) and labeled with ^{32}P by T4 polynucleotide kinase (16). Nick-translated probes were labeled with ^{32}P by *Escherichia coli* DNA polymerase I. Strand-specific probes made from single-stranded M13 templates were labeled with ^{32}P by *E. coli* DNA polymerase after hybridizing with hybridization probe primer (P-L Biochemicals).

Hybridizations with the oligonucleotide pools were carried out for 10–12 hr at 42°C in 0.45 M NaCl/45 mM sodium citrate/4× Denhardt's solution (17), 10% formamide, 0.1% $NaDodSO_4$, 50 mM Hepes (pH 7.0), and 200 μg of salmon testes DNA (Sigma) per ml. After hybridization, filters were washed twice in 0.3 M NaCl/30 mM sodium citrate/0.1% $NaDodSO_4$ for 20 min at 23°C and then washed once for 30 min at 42°C. Reduced-stringency hybridizations with nick-translated probes were carried out for 12–15 hr at 55°C in 0.9 M NaCl/90 mM sodium citrate/4× Denhardt's solution, 0.1% $NaDodSO_4$, 50 mM Hepes (pH 7.0), and 200 μg of salmon testes DNA per ml. High-stringency hybridizations with nick-translated and single-stranded probes were done for 18–24 hr under the same conditions as oligonucleotide hybridizations except for the formamide concentration, which was increased to 50%. Filters hybridized with nick-translated probes were washed under the same conditions as for oligonucleotide hybridizations except that the final washes were carried out at 50°C. Filters were stripped of probe for reuse in 50% formamide at 68°C for 1 hr.

DNA Sequence Analysis. DNA from λ clones was isolated (18) and characterized by restriction and hybridization analysis. DNA sequence analysis was done by the dideoxy chain-termination method (19) following subcloning into M13mp18 and M13mp19. In all cases, DNA sequence was determined for both strands.

*Present address: Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, CT 06457.
†Present address: Department of Virology, Chiron Corporation, Emeryville, CA 94608.

RESULTS

Protein kinases possess a 30-kDa domain responsible for their catalytic activity. Some sequences within the so-called kinase domain are universally conserved among all members of the family; others are conserved among only those members whose kinase activity is specific for tyrosine residues (2). Fig. 1 shows sequences conserved among protein kinases. The two underlined sequences have been found only among tyrosine-specific protein kinases and are highly conserved within this group (2).

In an effort to isolate tyrosine-specific protein kinase genes from yeast, mixed oligonucleotide pools encoding the two underlined sequences in Fig. 1 were designed as hybridization probes. Considering the complexity of the yeast genome and the requisite degeneracy of the probes, the optimal length of these oligonucleotides was calculated to be 17 bases. The probability of random occurrences of these sequences within the yeast genome was low enough so as not to be a major concern. Results discussed in this study will be limited to those obtained using oligonucleotides encoding the 6-amino-acid sequence: Asp-Val-Trp-Ser-Phe-Gly (DVWSFG), since oligonucleotides encoding the other conserved sequence in Fig. 1 failed to detect sequences within any protein kinase genes (data not shown).

Cloning *KIN1* from *S. cerevisiae*. The oligonucleotide pools shown in Fig. 2 were used to probe genomic yeast DNA at high stringency. These pools include all of the sequences that encode the sequence DVWSFG. Since there are six serine codons, it was necessary to use two oligonucleotide pools. DNA fragments were separated by agarose gel electrophoresis after digestion with *Eco*RI and transferred to nitrocellulose filters for hybridization. The yeast genome contained no sequences homologous to members of pool 2, but members of pool 1 hybridized to two DNA fragments with lengths of 5 kb and 10 kb (Fig. 2).

Oligonucleotide pool 1 was used to screen a library of yeast genomic DNA in bacteriophage λ. Positive signals were detected from 15 of ≈7000 plaques screened. DNA from each positive clone was isolated and subjected to restriction and hybridization analysis (data not shown). Among these clones, 14 contained sequences from the 5-kb *Eco*RI fragment that hybridized to the oligonucleotide probe. The remaining clone contained sequences from the 10-kb fragment. The sequences responsible for hybridization resided on a 750-bp *Sau*3AI fragment within the 5-kb *Eco*RI fragment and a 1.1-kb *Sph* I–*Kpn* I fragment within the 10-kb *Eco*RI fragment. These fragments were subcloned into M13 and sequenced. The 750-bp fragment contained the sequence: 5′ GATGTTTG-GTCGTTCTGG 3′, which was identical to one member of the probe pool, except for an inserted thymine residue (underlined). Since the reading frame defined by this sequence was closed 30 codons N-terminal- and 5 codons C-terminal to the region homologous to the oligonucleotide probe, and the subclone possessed no additional homology to protein kinase genes, no further work was done with this cloned sequence.

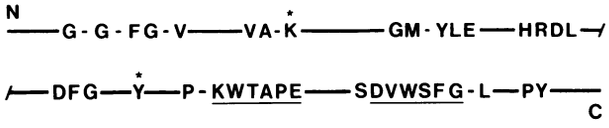

FIG. 1. Amino acids conserved among tyrosine-specific protein kinases. The underlined sequences are especially diagnostic of tyrosine kinases, but some residues are also conserved among serine/threonine-specific protein kinases. The first asterisk indicates the highly conserved lysine (K) residue within the ATP binding domain; the second indicates the phosphorylated tyrosine (Y) residue conserved among all known tyrosine kinases.

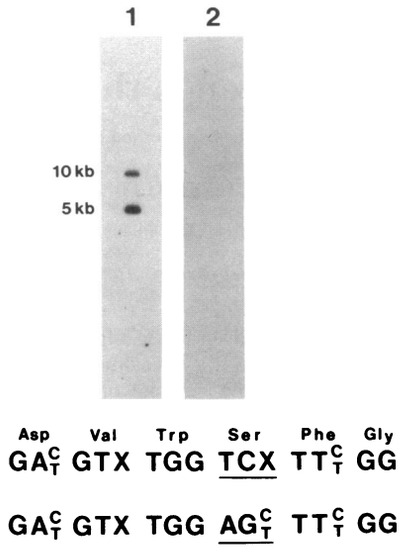

FIG. 2. Hybridization of degenerate oligonucleotide probes to genomic yeast DNA. DNA was digested with *Eco*RI and after electrophoresis through 1% agarose (5 μg per lane) was transferred to nitrocellulose. Oligonucleotides were pools of 17-mers encoding the 6-amino acid sequence DVWSFG. Since there are six serine codons, it was necessary to synthesize two pools differing only at the serine codons. Pool 1 (lane 1) contained a mixture of 64 probes, whereas pool 2 (lane 2) was 32-fold degenerate. The film was exposed for 18 hr at −70°C with an intensifying screen.

The 1.1-kb *Sph* I–*Kpn* I fragment possessed the sequence: 5′ GATGTCTGGTCATTTGG 3′, a perfect match to one member of the probe pool. The reading frame defined by this homology was open throughout the entire subclone. In addition, this reading frame contained the other hallmark sequences indicative of protein kinases (see below). This gene was designated *KIN1* (for kinase).

Cloning *KIN2*. A DNA fragment containing *KIN1* sequences was used as a hybridization probe against yeast genomic DNA under conditions of reduced stringency to identify *KIN1*-related sequences. Fig. 3 shows that, in addition to *KIN1*, one other DNA species hybridized to a 1.7-kb *Eco*RV–*Kpn* I fragment encoding the putative kinase domain of *KIN1*. This *KIN1* probe was used to screen the same library filters used for the initial oligonucleotide screen following denaturation to remove the old probe. Positive signals were obtained from six additional plaques under

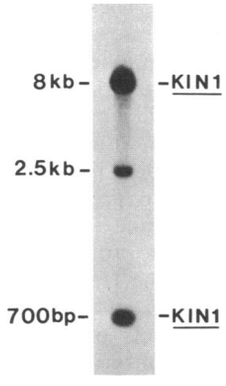

FIG. 3. Hybridization of a *KIN1* probe to genomic yeast DNA under conditions of reduced stringency. DNA (5 μg) was digested with *Hin*dIII. A 1.7-kb *Eco*RV–*Kpn* I fragment encoding the kinase domain of *KIN1* was used as probe. The 8-kb and 700-bp hybridizing fragments are from *KIN1*. The film was exposed for 18 hr at −70°C with an intensifying screen.

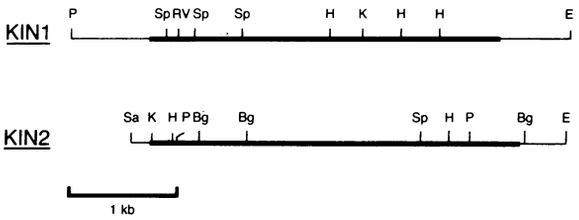

FIG. 4. Restriction maps of *KIN1* and *KIN2*. Both genes are 5′ → 3′ as drawn, with the heavy lines denoting coding sequence. P, *Pst* I; Sp, *Sph* I; RV, *Eco*RV; H, *Hin*dIII; E, *Eco*RI; K, *Kpn* I; Sa, *Sac* I; and Bg, *Bgl* II.

conditions of reduced stringency. Four of these clones contained a 2.5-kb *Hin*dIII fragment that hybridized to the *KIN1* probe only at low stringency. The remaining clones contained *KIN1* sequences, as judged by restriction analysis and high-stringency hybridization to the *KIN1* probe (data not shown). The 2.5-kb *Hin*dIII fragment was subcloned and subjected to DNA sequence analysis. This fragment possesses an open reading frame that spans the entire length of the fragment and shows homology to protein kinases and extensive homology to *KIN1*. The gene was therefore designated *KIN2*. Restriction maps of *KIN1* and *KIN2* are shown in Fig. 4. Probes derived from *KIN2* sequences failed to detect additional members of this gene family within the yeast genome under hybridization conditions of reduced stringency (data not shown).

Expression of *KIN1* and *KIN2*. Single-stranded probes predicted to detect *KIN1* or *KIN2* mRNA were hybridized with yeast poly(A)$^+$ RNA. Fig. 5 shows that both genes are transcribed into 3.5-kb mRNAs in proliferating a cells.

***KIN1* and *KIN2* Encode Protein Kinases.** Complete DNA sequence analysis of the *KIN* genes revealed that *KIN1* encodes a 3.2-kb uninterrupted open reading frame corresponding to a polypeptide of predicted size 1064 amino acids (117 kDa). *KIN2* encodes a 3.4-kb uninterrupted open reading frame corresponding to a polypeptide of predicted size 1152 amino acids (126 kDa). These values assume the use of the 5′-most methionine codons in the open reading frames. No consensus sequences for intron splicing (20) were identified in either gene, or in sequences 5′ (1 kb for *KIN1* and 300 bp for *KIN2*) to the predicted initiation sites. An amino acid

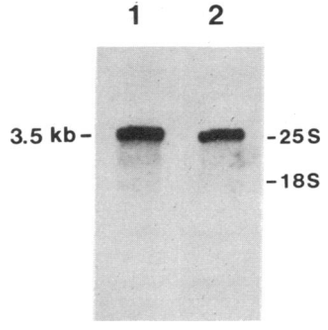

FIG. 5. Analysis of yeast RNA by hybridization to strand-specific probes of *KIN1* and *KIN2*. Ten micrograms of poly(A)$^+$ RNA was denatured, electrophoresed through a 6% formaldehyde/1.2% agarose gel, and transferred to nitrocellulose. The probe for the *KIN1* message (lane 1) was derived from the 1.1-kb *Sph* I–*Kpn* I fragment cloned into M13mp18. The probe for *KIN2* (lane 2) was derived from the 2.4-kb *Hin*dIII–*Sph* I fragment cloned into M13mp18. After hybridizing the hybridization probe primer to single-stranded template DNA from these clones, ^{32}P was incorporated by DNA synthesis away from the cloned sequences. Markers are the 18S (1.7 kb) and 25S (3.3 kb) rRNAs visualized after staining with ethidium bromide. Films were exposed for 14 hr at −70°C with an intensifying screen.

alignment of *KIN1* and *KIN2* is shown in Fig. 6. Both genes carry their kinase domains near the N terminus of the predicted proteins. Throughout this region these proteins are 90% identical. Outside of the kinase domain, homology is patchy, with some regions highly conserved and others quite thoroughly diverged.

The putative kinase domains of *KIN1* and *KIN2* were compared to sequences in the Genbank protein database.‡ Although homology was identified between the new loci and many protein kinases, the yeast genes did not correspond to previously isolated protein kinase genes from any species. The two proteins most closely related to these unusual yeast proteins were the bovine cAMP-dependent protein kinase and the protein kinase encoded by the v-*src* oncogene (21, 22). Fig. 7 shows that the kinase domain of *KIN1* and *KIN2* is 27% and 25% identical to that of v-*src*, respectively, and 30% identical to that of bovine cAMP-dependent protein kinase. Sequences C-terminal to the kinase domains of *KIN1* and *KIN2* showed no homology to any sequence in the Genbank protein database. The DNA sequences of these genes are available through the Genbank nucleic acid database.§

DISCUSSION

Isolation of *KIN1* and *KIN2*. We report the use of mixed oligonucleotide pools encoding conserved regions of tyrosine-specific protein kinases to probe the genome of budding yeast. These probes allowed the isolation of two yeast genes that possess homology to protein kinases. These genes are very closely related to each other (see Fig. 7) but were not identifiable as protein kinase genes previously isolated from any species and, therefore, were designated *KIN1* and *KIN2* (for kinase). *KIN1* was detected by its hybridization to one of the oligonucleotide probes. *KIN2* failed to hybridize to these oligonucleotides and was isolated using *KIN1* sequences as probes under conditions of reduced stringency. *KIN2* possesses a sequence within its kinase domain that is homologous, except for a single mismatched base, to a member of the oligonucleotide pool used to isolate *KIN1*. This mismatch resulted from a valine to isoleucine change in the target sequence, DVWSFG, suggesting that the conditions used for oligonucleotide hybridization were relatively stringent. Considering the paucity of yeast genes that possess the target sequence, our results suggest that yeast is not likely to possess a large number of genes encoding tyrosine-specific protein kinases unless the catalytic domains of these enzymes are substantially different in structure from those of metazoan tyrosine kinases.

***KIN1* and *KIN2* Encode Unusual Protein Kinases.** *KIN1* and *KIN2* encode predicted polypeptides of 1064 amino acids and 1148 amino acids, respectively. An unusual feature of these genes is that their putative kinase domains reside near the N termini of relatively large proteins. Most protein kinases carry their catalytic domains at or near their C termini. The only reported precedent for a protein kinase with an N-terminal kinase domain and extensive C-terminal sequence is the product of the v-*abl* oncogene (23). A hydropathy plot indicated that neither of the yeast genes contains a sequence of sufficient length and hydrophobicity to constitute a potential transmembrane domain (data not shown), similar to those found in protein kinases that act as growth factor receptors (24). This finding, however, does not preclude a

‡EMBL/Genbank Genetic Sequence Database (1986) Genbank (Bolt, Beranek, and Newman Laboratories, Cambridge, MA), Tape Release 40.

§EMBL/Genbank Genetic Sequence Database (1987) Genbank (Bolt, Beranek, and Newman Laboratories, Cambridge, MA), Tape Release 48.

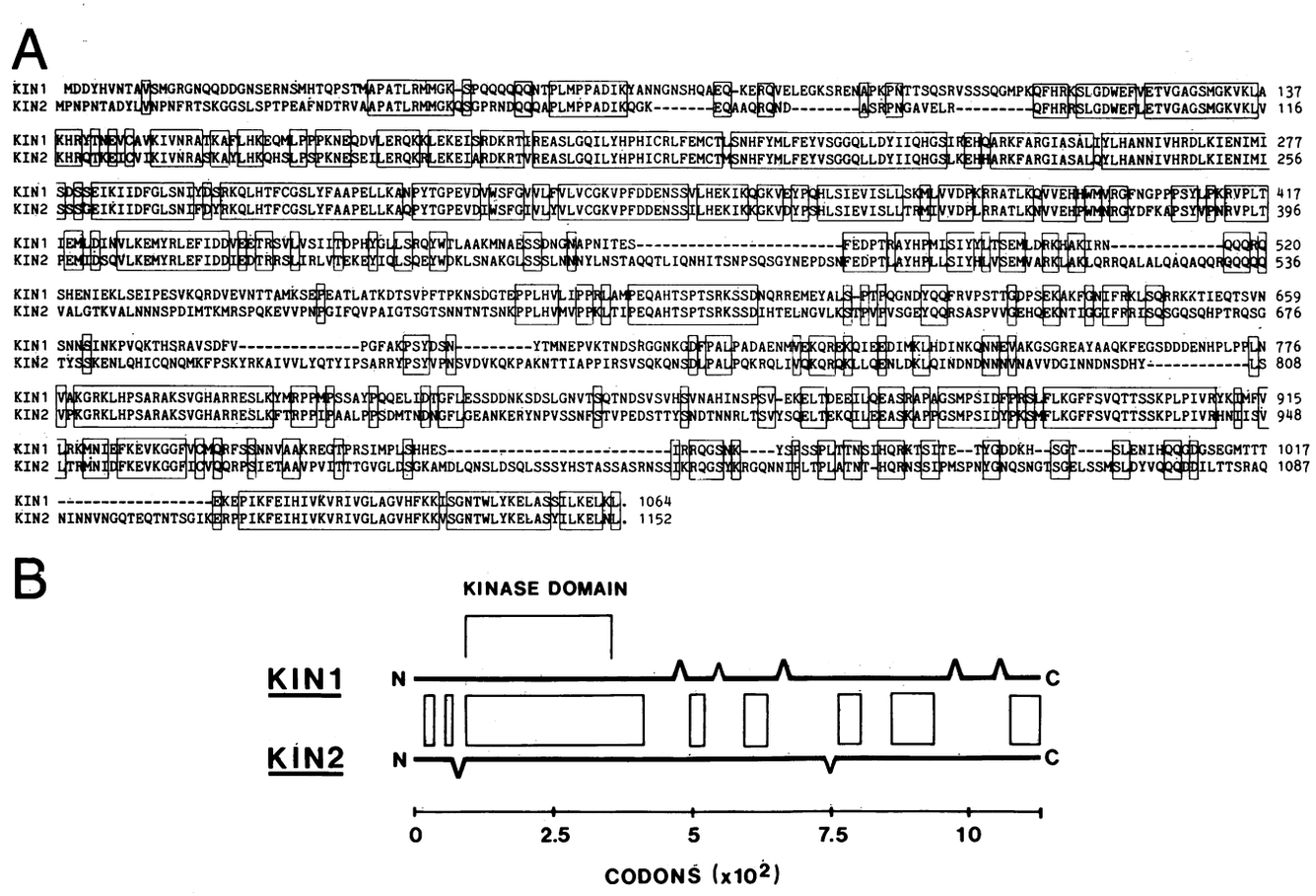

FIG. 6. Predicted gene products of *KIN1* and *KIN2*. (*A*) Alignment of amino acid sequences deduced from *KIN1* and *KIN2*. In both cases, the first methionine codon within the open reading frame was used to define the N terminus. Gaps were introduced as indicated by dashes. (*B*) Topographical alignment of the *KIN* proteins. Boxed areas indicate extended regions of homology (>75% amino acid identities).

weaker association with the membrane, as in the case of the protein encoded by v-*fps* (25).

The kinase domains of *KIN1* and *KIN2* are unusual because they possess some structural features that are distinctive of serine/threonine-specific protein kinases and other features that are characteristic of tyrosine-specific kinases. A protein homology search indicated that the two protein kinases most closely related to *KIN1* and *KIN2* are (*i*)

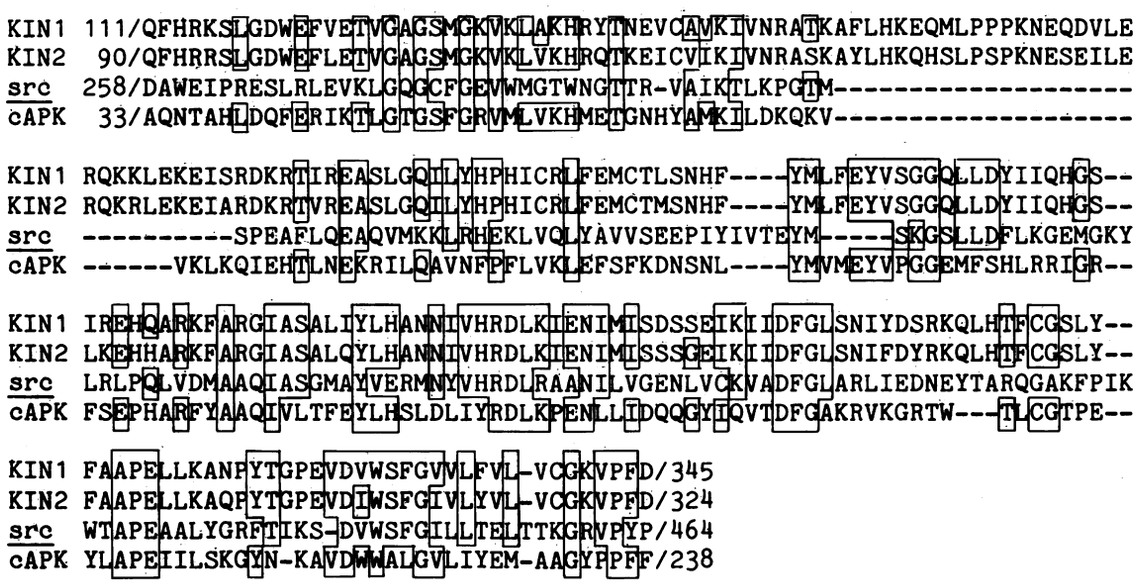

FIG. 7. Alignment of amino acid sequences deduced from *KIN1* and *KIN2* catalytic domains with those of bovine cAMP-dependent protein kinase and v-*src*. The deduced amino acid sequences for bovine cAMP-dependent protein kinase (cAPK) and v-*src* have been published (21, 22). The sequences were aligned by eye for maximum homology and gaps were introduced as indicated by dashes. Identities exclusively between the *KIN* sequences are not boxed.

the bovine cAMP-dependent protein kinase, a serine/threonine-specific enzyme, and (*ii*) the product of the v-*src* oncogene, the prototypical tyrosine-specific protein kinase. Notably absent from the yeast genes is the major phosphorylated tyrosine residue (Tyr-416) in the v-*src* product, which is conserved among all known tyrosine kinases (2). Biochemical analysis of the *KIN1* and *KIN2* gene products will allow a determination of the substrate specificity of these unusual enzymes. This will undoubtedly lead to refinements in the diagnostic features of both subfamilies of protein kinases.

The Physiological Functions of *KIN1* and *KIN2* Are Not Known. *KIN1* and *KIN2* are transcribed into 3.5-kb mRNAs in proliferating yeast cells, but we presently have no hint as to the functions of these genes. In an effort to define the functions of *KIN1* and *KIN2*, we have constructed deletion mutants of both genes (unpublished). Cells lacking *KIN1*, *KIN2*, or both genes are viable and appear to grow normally under a variety of conditions. Thus far, we have been unable to detect any phenotypic defects associated with null mutants of the *KIN* genes. We have therefore extended our study of *KIN* genes to include the fission yeast *Schizosaccharomyces pombe*. Results of this work will be reported elsewhere.

We thank I. Herskowitz, S. Michaelis, A. Mitchell, and P. Sternberg for helpful discussions. This work was supported by grants from the National Institutes of Health (Grants 08CA012705 and S07RR05355) and by funds from the G. W. Hooper Research Foundation. D.E.L. was supported by a postdoctoral fellowship from the Damon Runyon–Walter Winchell Cancer Fund (Fellowship DRG790). Computer resources were provided by BIONET National Computer Resource for Molecular Biology.

1. Cohen, P. (1982) *Nature (London)* **296,** 613–620.
2. Hunter, T. & Cooper, J. A. (1985) *Annu. Rev. Biochem.* **54,** 897–930.
3. Bishop, J. M. & Varmus, H. (1985) in *RNA Tumor Viruses*, ed. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 999–1108.
4. Lorincz, A. T. & Reed, S. I. (1984) *Nature (London)* **307,** 183–185.
5. Patterson, M., Sclafani, R. A., Fangman, W. L. & Rosamond, J. (1986) *Mol. Cell. Biol.* **6,** 1590–1598.
6. Celenza, J. L. & Carlson, M. (1986) *Science* **223,** 1175–1180.
7. Teague, M. A., Chaleff, D. T. & Errede, B. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 7371–7375.
8. Simon, M., Seraphin, B. & Faye, G. (1986) *EMBO J.* **5,** 2697–2701.
9. Schieven, G., Thorner, J. & Martin, G. S. (1986) *Science* **231,** 390–393.
10. Petes, T. D. (1980) *Annu. Rev. Biochem.* **49,** 845–876.
11. Winston, F., Chumley, F. & Fink, G. R. (1983) *Methods Enzymol.* **101,** 211–228.
12. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
13. Jensen, R., Sprague, G. F. & Herskowitz, I. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3035–3039.
14. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–183.
15. Sinha, N. D., Biernat, J., McManus, J. & Koster, H. (1984) *Nucleic Acids Res.* **12,** 4539–4557.
16. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
17. Jay, E., Bambara, R., Padmanabhan, R. & Wu, R. (1974) *Nucleic Acids Res.* **1,** 331–354.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), p. 85.
19. Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* **94,** 441–448.
20. Langford, C. J. & Gallwitz, D. (1983) *Cell* **33,** 519–527.
21. Shoji, S., Parmelee, D. C., Wade, R. D., Kumar, S., Ericsson, L. H., Walsh, K. A., Neurath, H., Long, G. L., Demaille, J. G., Fischer, E. H. & Titani, K. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 848–851.
22. Czernilofsky, A. P., Levinson, A. D., Varmus, H. E., Bishop, J. M., Tischer, E. & Goodman, H. M. (1980) *Nature (London)* **287,** 198–203.
23. Reddy, E. P., Smith, M. J. & Srinivasan, A. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3623–3627.
24. Ullrich, A., Gray, A., Tam, A. W., Yang-Feng, T., Ysubokawa, M., Collins, C., Henzel, W., Le Bon, T., Kathuria, S., Chen, E., Jacobs, S., Francke, U., Ramachandran, J. & Fujita-Yamaguchi, Y. (1986) *EMBO J.* **5,** 2503–2512.
25. Woolford, J. & Beemon, K. (1984) *Virology* **135,** 168–180.