BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **245,** 177–184 (1998)
ARTICLE NO. RC988380

Expression Pattern of the *C. elegans* p21-Activated Protein Kinase, CePAK

Yuichi Iino[1] and Masayuki Yamamoto
Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, P.O. Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received February 25, 1998

The *C. elegans* p21-activated protein kinase (CePAK) has a high amino-acid sequence similarity to mammalian PAKs. Tissue specificity of the expression of CePAK was examined using *lacZ* and GFP reporters. This analysis indicated that CePAK is expressed mainly in pharyngeal muscles, the CAN neurons, and motor neurons in the ventral nerve cord, as well as several cells in the tail region and the distal tip cells. The CePAK::GFP fusion protein was preferentially localized to the cell surface in pharyngeal muscles. © 1998 Academic Press

p21-activated protein kinases (PAKs) constitute a family of protein kinases which are activated by the small GTPase binding proteins Cdc42 and Rac (1). One of their yeast homologues, Ste20p, is a prototype of this family and was identified as a protein kinase required to transmit the mating pheromone signal form $G_{\beta\gamma}$ to the downstream MAP kinase cascade (2). Later analysis suggested that Ste20p phosphorylates the MAPKKK Ste11p in this cascade (3). Another member of this family in yeast, Cla4p, has an overlapping function with Ste20p and is required for localized growth of the bud (4). Cdc42p is required for bud emergence, suggesting the functional link between these two proteins. The fission yeast homologue of PAKs, Shk1p/Pak1p, is implicated in polarized growth. Cells lacking shk1/pak1 arrests growth as small round cells (5, 6). Deletion of the fission yeast *cdc42* gene results in lethality with a morphology similar to that of the *shk1* mutant, whereas overexpression of the activated forms results in irregular cell morphology (7). Several subtypes are also found in mammalian PAKs, some of which have been shown to be bound and activated by Cdc42 and Rac (1). As in yeast, PAK has been shown to activate the MAP kinase JNK/SAPK (8). Recent report showed also that it regulates the reorganization of the actin cytoskeleton (9). Although these observations suggest the role for the PAK family in signal transduction and cell morphogenesis, little is known about their *in vivo* function in multicellular organisms. *C. elegans* is one of the ideal organisms to study gene function in multicellular context, because of the availability of powerful genetics (10) and its simple anatomy made of about 1000 identifiable cells (11). Here we report the expression pattern of CePAK, a *C. elegans* ortholog of PAK, revealed by reporter study. CePAK is expressed in pharyngeal muscles and several neurons, and is localized preferentially at the cell surface.

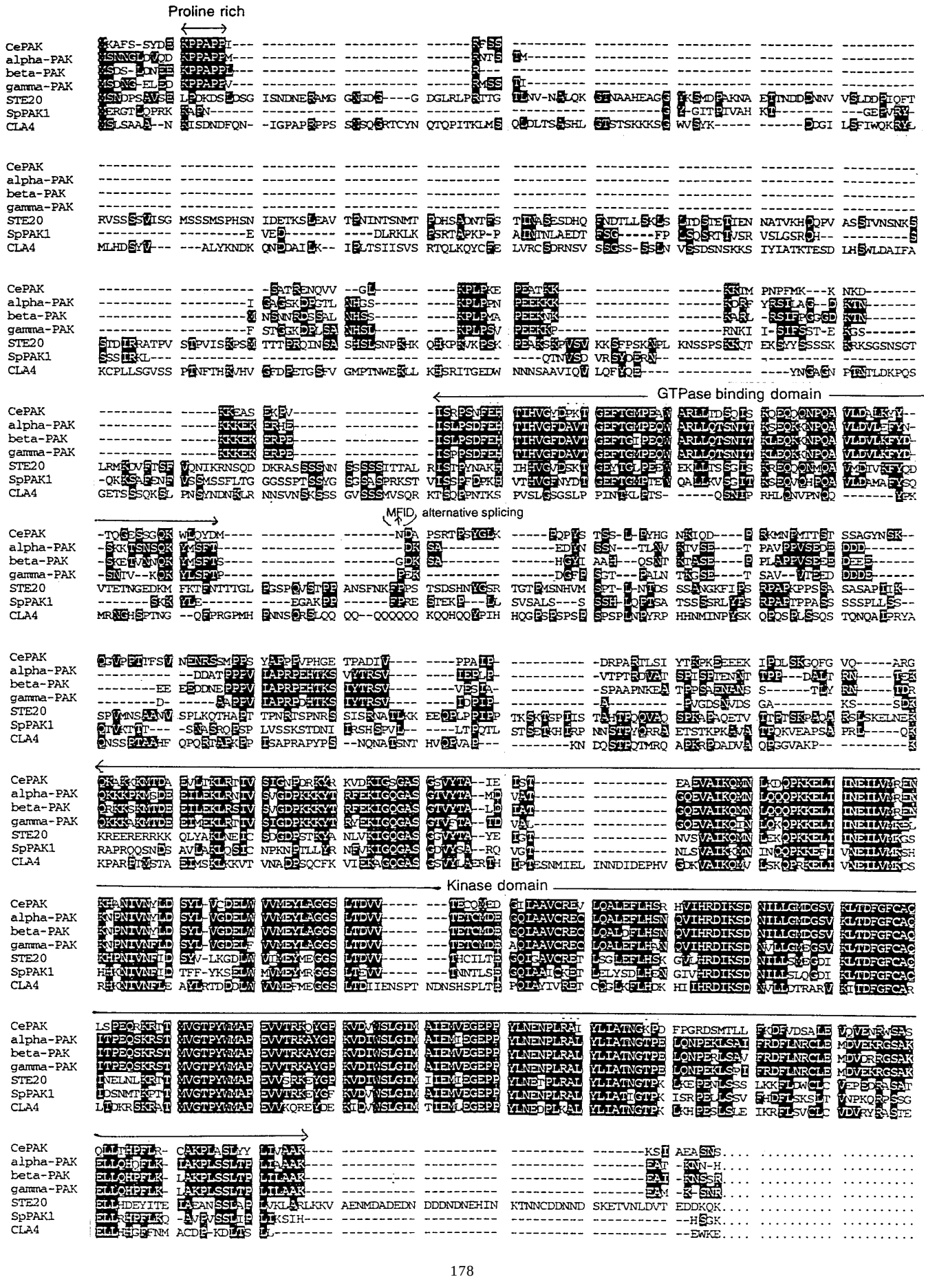
[1] Corresponding author: Fax: +81 3 5802 2042. E-mail: iino@ims.u-tokyo.ac.jp.

MATERIALS AND METHODS

Isolation of cDNAs. The original cDNA, yk116f6, was isolated and partially sequenced in the cDNA project led by Yuji Kohara (National Institute of Genetics, Mishima, Japan), and generously provided to us. The 5′ clones were isolated by PCR using the *C. elegans* cDNA library (12) as a template. The primers used were the antisense primer Preprimcepak (GTATGATGGCGTCCGAGAAGGT) corresponding to the sequence in yk116f6 and the primer Preprim (AGCAAAGCTTAAAGGAATTCGATT) corresponding to the vector sequence.

Construction of the reporter plasmids. A set of vectors harboring *lacZ*- or GFP- coding sequence with or without a nuclear localization signal (NLS) generated by A. Fire and colleagues (Carnegie Institution of Washington, Baltimore, USA) were used in this study (13). The genomic cosmid clone C45B2 was provided by A. Coulson (Sanger Center, UK). The ~11kb genomic fragment spanning the *Nhe*I and the *Bgl*II site was excised from the cosmid C45B2 as the promoter region. This fragment includes ~10kb upstream and ~1kb downstream of the initiation codon of CePAK. The *Bgl*II site is in the 4th exon. This fragment was subcloned between the *Pst*I and *Bam*HI sites of the pPD22.11 vector that has *lacZ* with NLS, to generate pCPK1.11. The same genomic fragment was subcloned between the *Pst*I and *Bam*HI sites of the pPD95.70 vector that has GFP with NLS, to generate pCPK1.70. For promoter fusion with the GFP reporter without NLS, a *Pst*I site was created just upstream of the initiation codon and the genomic fragment spanning the *Nhe*I site to this *Pst*I site was subcloned into the *Pst*I site of pPD95.81, the GFP vector without NLS, to generate pCPK1.81. For protein fusion, the genomic fragment spanning the *Nhe*I and *Bgl*II sites described above was fused to the cDNA fragment from the *Bgl*II site to the end of cDNA to generate pCPK2. pCPK2 was further subcloned by excising the fragment from the *Nhe*I site to another *Nhe*I site that is located close

0006-291X/98 $25.00


Proline rich
CePAK
alpha-PAK
beta-PAK
gamma-PAK
STE20
SpPAK1
CLA4
GTPase binding domain
MFID
alternative splicing
Kinase domain

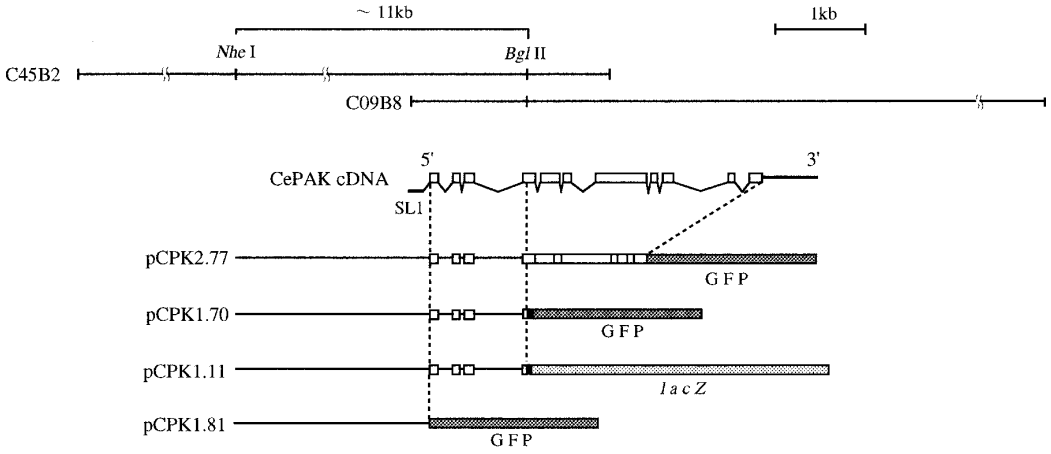


FIG. 2. Genomic structure of the CePAK mRNA and structure of the reporter constructs. The splicing pattern for the CePAK mRNA deduced from the cDNA sequence compared with that of the cosmids C45B2 and C09B8 is shown. The structure of the reporter constructs is also depicted. Open boxes indicate the CEPAK coding sequence. Hatched boxes indicate coding sequence for GFP and *lacZ*. Filled boxes indicate the nuclear localization signal. All the plasmids had the 3′ region of *unc-54* placed downstream of the coding sequences.

to the end of the ORF and subcloning into the *Nhe*I-*Kpn*I sites of pPD95.77, the GFP vector without NLS, to generate pCPK2.77.

Germline transformation and staining. The reporter plasmids were coinjected individually or in combination with pNH86 harboring the wild-type *dpy-20* gene into the syncytial gonad of the *dpy-20 (e2017)* hermaphrodite adults by the standard microinjection technique (14). Non-Dpy transformants that heritably transmit the transgene to the progeny were selected and observed as below. Injected plasmids form a tandem array and extrachromosomally maintained and transmitted in an unstable fashion. Note therefore that mosaic expression pattern is often observed for transgenes.

For histochemical detection of the β-galactosidase activity, animals were fixed and stained with X-Gal as described (13). For detection of GFP, intact animals were observed under the fluorescent microscope Axiophot (Zeiss) and recorded using chilled a CCD camera and image acquisition system (Argus 50, Hamamatsu Photonics).

RESULTS

Structure of the CePAK cNAs

C. elegans cDNA project by Yuji Kohara and his colleagues identified a group of cDNAs (CELK02351) whose partial sequence predicted a protein related to mammalian p21-activated protein kinases (PAKs). We determined the complete nucleotide sequence of the longest cDNA in this group, yk116f6. Because yk116f6 seemed to lack the 5′ end, this region was isolated from the cDNA library by PCR. Two of the five 5′ clones thus isolated and sequenced had the sequence of SL1, the trans-spliced leader found commonly at the 5′ end of *C. elegans* mRNA molecules. Therefore, these clones were supposed to correspond to the 5′ end of the mRNA. The assembled cDNA sequence was deposited to DDBJ under the accession number D83215. The amino acid sequence predicted from these cDNAs had a high similarity to mammalian PAK proteins (Fig. 1). It also had similarity to *S. cerevisiae* Ste20p, Cla4p and *S. pombe* Pak1p/Shk1p, but it was more similar to mammalian PAKs. Comparison of the cDNA sequence with the nucleotide sequence reported by the *C. elegans* genome sequencing project indicated that this gene is located on the cosmids C09B8 and C45B2, which are mapped to the center of chromosome X. This comparison also revealed the existence of 10 introns in the CePAK gene (Fig. 2). One of the five 5′ clones had a 9-nucleotide insertion compared to others apparently resulting from an alternative utilization of the splicing acceptor site for intron 4. N^{139} just after the GTPase-binding domain was substituted by four amino acids MFID in this clone (Fig. 1).

Expression Pattern of CePAK

To determine the tissue-specific expression of CePAK, reporter plasmids were constructed using standard *C. elegans* vectors, in which about 10kb of the genomic sequence upstream of the initiation codon of CePAK was tagged to *lacZ* with a nuclear-localization signal (NLS), or GFP with or without NLS. These plasmids were introduced into *C. elegans* hermaphrodites and localization of β-galactosidase or GFP was determined.

Prominent expression of CePAK was detected in the pharynx and two laterally located cells in L1 larva, as

FIG. 1. Comparison of the amino acid sequence of CePAK with that of other PAK family members. Amino acid sequence deduced from the CePAK cDNAs was compared with that of human αPAK, βPAK, γPAK, *S. cerevisiae* STE20p, *S. pombe* PAK1p and *S. cerevisiae* CLA4p. The proline-rich sequence near the N-terminus in CePAK and human PAKs, the GTPase binding domain (CRIB), and the kinase domain are indicated. The sequence alteration resulting from alternative splicing was also shown.

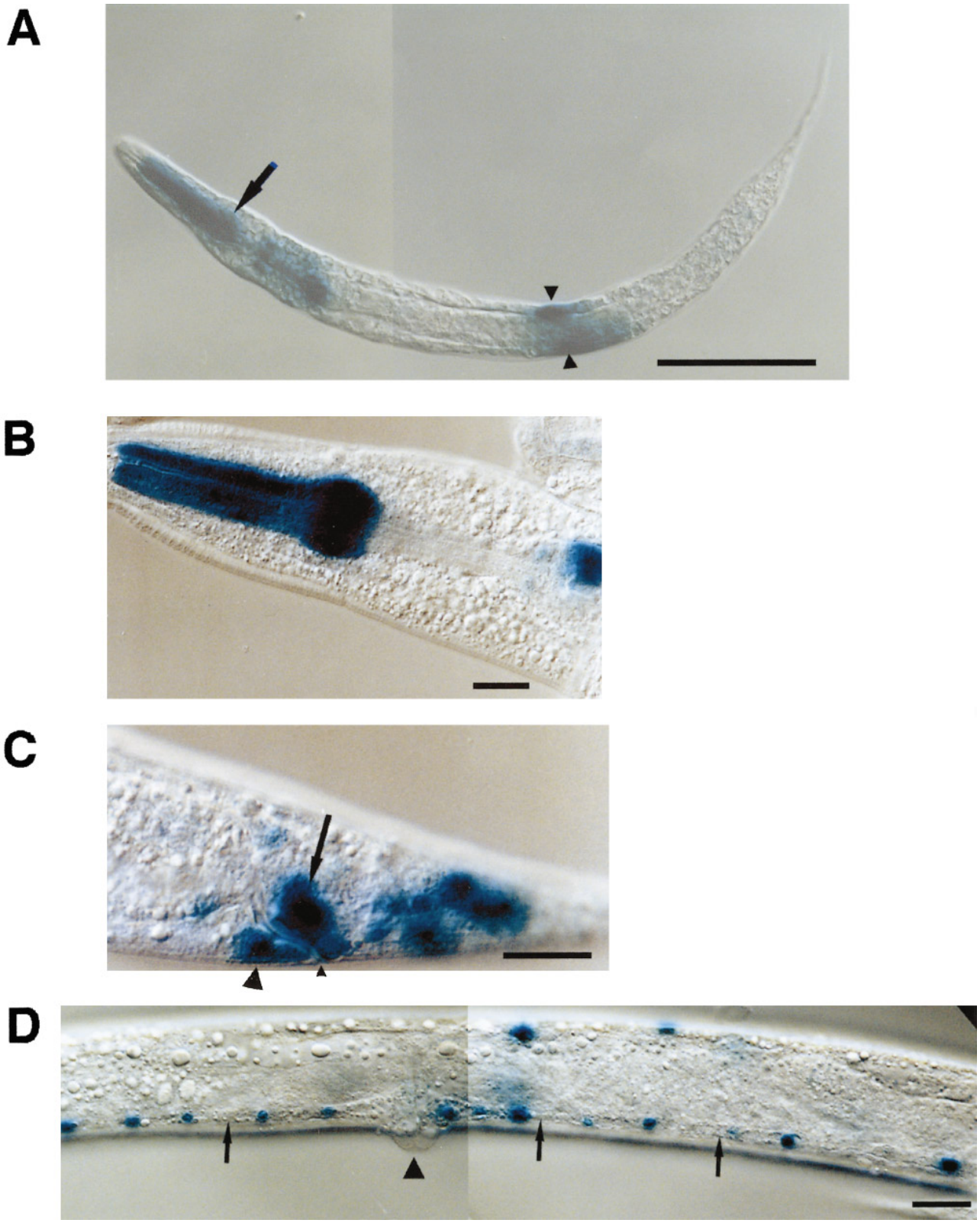


FIG. 3. Expression pattern of the CePAK-NLS-*lacZ* reporter. pCPK1.11 (Fig. 2) was introduced into the *dpy-20* strain, and the transformed animals were fixed and stained for β-galactosidase. A. An L1 animal showing expression in the pharynx (arrow) and the CAN neurons (arrowheads). B. Staining of the pharynx in an adult. Although the reporter had a nuclear localization signal, not only the nuclei but also the whole muscle cells were stained because of strong expression and leakage out of the nuclei. C. Staining in the tail region. Nuclei of the B cell (arrow) and the Y cell (large arrowhead) and several unidentified cells were stained in an adult. The small arrowhead indicates the anus. D. Staining in ventral cord neurons. A row of neuronal nuclei in the ventral nerve cord (arrows) are stained. The vulva is indicated by the arrowhead. Bar: 20μm.

revealed by staining of the NLS-*lacZ* reporter strain (Fig. 3A, B). The pharyngeal cells are muscle cells which occupy most of the body of pharynx and the lateral cells were identified as the CAN neurons based on the position of the nucleus and axon morphology (Fig. 5, E-H). Both pharyngeal muscle cells and CAN neurons start to express the reporter as embryo and do so until adulthood. Also stained were several cells in the tail region (Fig. 3C). These include the B and Y cells from L1 to adult, the hypodermal blast cell T in the L1 and some of its progeny in later stages. The B cell and the Y cell are male-specific blast cells that give rise to male-specific structures such as spicules. In hermaphrodites, these cells serve simply as support cells in L1, but Y becomes a motor neuron later. A subset of cells in the ventral nerve cord were also stained (Fig. 3D). These were identified as DA and DB neurons by GFP observation of intact L1 animals (data not shown). These neurons are cholinergic motor neurons innervating dorsal body wall muscles required for locomotion.

The observation of the transformants with the GFP reporter without NLS further revealed expression of CePAK in the distal tip cell (DTC, Fig. 4A, B), the spermatheca (Fig. 4C, D), and the uterus (Fig. 4E, F). DTC migrates and leads the expansion of the gonad during the larval development in hermaphrodites. It also regulates the mitosis-to-meiosis transition of germline nuclei in the gonad. The spermatheca is the compartment in the gonad where mature sperm is stored and fertilization occurs when oocytes are ovulated from the ovary into the uterus. It was not clear whether GFP was expressed by the somatic cells of the gonad or by the germline cells, namely sperm, in the spermatheca. The expression in the uterus was seen outside the embryos. Again, it was not clear which cells expressed the reporter in the uterus, although observation of NLS-*lacZ* and NLS-GFP indicated that some somatic cells of the gonad expressed CePAK (data not shown).

Observation of the embryos expressing the GFP reporter without NLS revealed that CePAK is expressed in many hypodermal cells surrounding the embryos during the morphogenic stage (Fig. 4G-J).

Intracellular Localization of CePAK

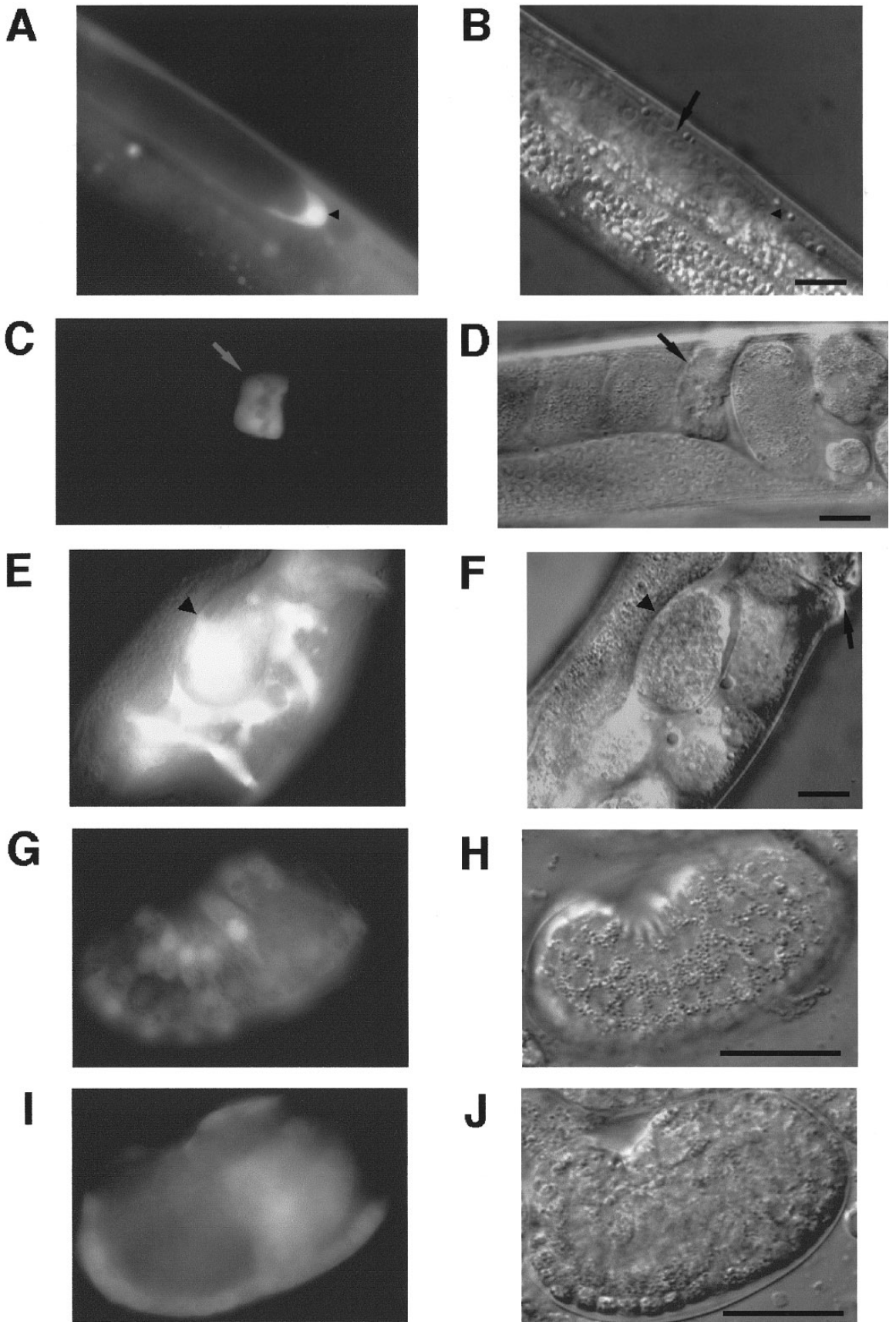
To determine the intracellular localization of the CePAK protein, the C-terminal two amino acids of CePAK were replaced with the GFP coding sequence, and the fusion protein was expressed from the CePAK promoter used in the above analysis (Fig. 2, pCPK2.77). Compared with the control in which only GFP was expressed (Fig. 5C, D), CePAK::GFP was enriched at the cell surface in pharyngeal muscle cells. It was especially condensed at cell boundaries (Fig. 5A, B). In the CAN neurons, CePAK::GFP was detected in the cell bodies and along the axons (Fig. 5E, F). It was not clear because of the small size of these structures whether CePAK::GFP is localized at the cell surface or within the cytoplasm. The localization of the protein in other cell types was not determined because of the weak fluorescence.

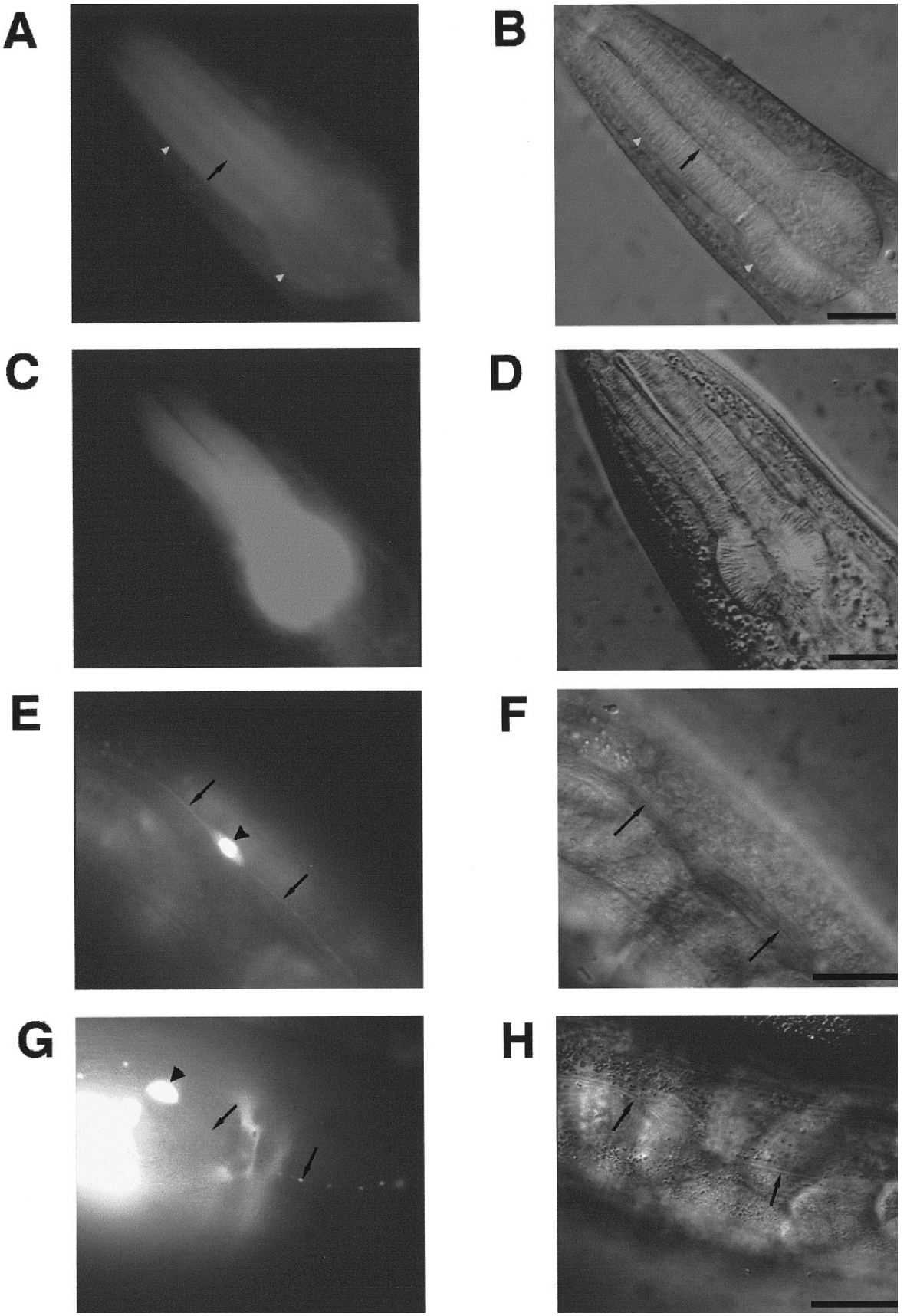
DISCUSSION

We have examined the expression pattern of CePAK, *C. elegans* p21-activated protein kinase, by using transgenic strains harboring the reporter genes driven by the promoter region of CePAK. Although we have no evidence that the expression of the reporters we employed reflects precisely the localization of the authentic CePAK mRNA, it is likely because transcriptional regulatory elements are generally located in a relatively small region upstream of transcription start sites in *C. elegans*, and our reporter plasmids include sufficiently large upstream regions (~10kb). Our analysis showed that CePAK is expressed in a small number of cells. The CAN neuron, where very strong expression of CePAK was detected, has a characteristic structure. It extends its axons along the excretory canal, a structure implicated in osmoregulation, in close association with it. CAN forms few, if any, synapses to other neurons and a few possible gap junction to the excretory canal (15). A laser ablation experiment showed that this neuron is essential for survival (15). These observations led to the assumption that CAN regulates the function of the ex-

FIG. 4. Expression pattern of the CePAK-GFP reporter without NLS. pCPK1.81 (Fig. 2) was introduced into the *dpy-20* strain and the transformed animals were observed. Each pair of panels show the fluorescent image (A, C, E, G and I) and the Nomarski image (B, D, F, H and J) viewed at the same focal plane. A, B. Expression in the distal tip cell (DTC, arrowheads). The gonad (arrow in B) follows the DTC. Germ-line nuclei in the gonad are evident in B. C, D. Expression in the spermatheca (arrows). E, F. Expression in the uterus. Expression in one of the embryos is also seen (arrowhead). The vulva is shown by the arrow in F. G-J. Expression in the embryo. G, H and I, J show different focal planes of the same embryo. G, H. Close to the surface of the embryo. I, J. Along the midline of the embryo. Bar: 20μm.

FIG. 5. Localization of the CEPAK::GFP fusion protein compared with that of GFP. A, B, E and F. The CEPAK::GFP reporter pCPK2.77 was coinjected with the NLS-GFP reporter pCPK1.70 to mark the nucleus, and the transformed animals were observed. C, D, G and H. Transformants with the GFP reporter without NLS (pCPK1.81) were observed for comparison. A, C, E and G show the fluorescent images and B, D, F and H show the Nomarski images viewed at the same focal planes. A-D Pharynx in the adult. CePAK::GFP is preferentially localized at cell boudaries (arrows) and cell surface (arrowheads) in A. E-H. The CAN neuron in the adult. CePAK::GFP and GFP are both localized to cell bodies (arrowheads in E and G) and their processes (arrows in E and G) running along the excretory canal (arrows in F and H) on the lateral surface of adults. Bar: 20μm.


A
B
C
D
E
F
G
H
I
J


A
B
C
D
E
F
G
H

cretory canal. Ventral cord neurons DA and DB also expressed CePAK. These neurons are cholinergic motor neurons required for locomotion. DA and DB are supposed to act in reverse and forward movement, respectively (16). Although cell bodies of these neurons are located in the ventral cord, their axons extend circumferentially and finally innervate dorsal body wall muscles. Based on the known characteristics of PAKs and related kinases in various organisms, CePAK may act in these neurons to transmit signals regulating neuronal activity, or growth cone migration during development. Expression in the distal tip cells is intriguing in the latter view, because these cells are also migratory. Several mutations disrupting migration of both DTC and neuronal axons are known (17).

Examination of intracellular localization of the CePAK protein using the GFP tag showed that the protein is located at cell surface, especially at cell boundaries, in the pharyngeal muscles. Yeast Ste20p is localized at the growing tip of cell surface and this localization is dependent on its interaction with Cdc42p (18, 19). A more recent report showed that Ste20p and PAK interact with the G protein β subunit, which may also contribute to anchorage of the protein to the membrane through Gγ (20). The amino acid sequence at the C-terminal of the kinase domain of Ste20p and PAK that is required for interaction with Gβ is partially conserved in CePAK. PAKs are also reported to interact with the SH3 domain of Nck through the N-terminal proline-rich sequence (21). The proline-rich region is perfectly conserved in CePAK. A similar interaction may also contribute to membrane localization and activation of CePAK in *C. elegans*. Further study, including the isolation of loss of function and gain of function mutants, will elucidate the *in vivo* role of the CePAK protein and its interaction with other proteins in the cells identified in this study.

ACKNOWLEDGMENTS

We thank Yuji Kohara for cDNA clones, Alan Coulson for cosmids, and Andy Fire for vectors. The nematode strain used in this study was provided by the Caenorhabditis Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR).

REFERENCES

1. Lim, L., Manser, E., Leung, T., and Hall, C. (1996) *Eur. J. Biochem.* **242,** 171–185.
2. Leberer, E., Dignard, D., Harcus, D., Thomas, D., and Whiteway, M. (1992) *EMBO J.* **11,** 4815–4824.
3. Wu, C., Whiteway, M., Thomas, D., and Leberer, E. (1995) *J. Biol. Chem.* **270,** 15984–15992.
4. Cvrckova, F., De Virgilio, C., Manser, E., Pringle, J., and Nasmyth, K. (1995) *Genes Dev.* **9,** 1817–1830.
5. Marcus, S., Polverino, A., Chang, E., Robbins, D., Cobb, M., and Wigler, M. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6180–6184.
6. Ottilie, S., Miller, P., Johnson, D., Creasy, C., Sells, M., Bagrodia, S., Forsburg, S., and Chernoff, J. (1995) *EMBO J.* **14,** 5908–5919.
7. Miller, P., and Johnson, D. (1994) *Mol. Cell Biol.* **14,** 1075–1083.
8. Brown, J., Stowers, L., Baer, M., Trejo, J., Coughlin, S., and Chant, J. (1996) *Curr. Biol.* **6,** 598–605.
9. Sells, M., Knaus, U., Bagrodia, S., Ambrose, D., Bokoch, G., and Chernoff, J. (1997) *Curr. Biol.* **7,** 202–210.
10. Brenner, S. (1974) *Genetics* **77,** 71–94.
11. Sulston, J., and Horvitz, H. (1977) *Dev. Biol.* **56,** 110–156.
12. Hayashizaki, S., Iino, Y., and Yamamoto, M. (1998) *Genes to Cells* **3,** *in press.*
13. Fire, A., Harrison, S., and Dixon, D. (1990) *Gene* **93,** 189–198.
14. Mello, C., Kramer, J., Stinchcomb, D., and Ambros, V. (1991) *EMBO J.* **10,** 3959–3970.
15. White, J., Southgate, E., Thomson, J., and Brenner, S. (1986) *Philosophical Transactions of the Royal Society of London* **314B,** 1–340.
16. Chalfie, M., and White, J. (1988) *in* The nematode *Caenorhabditis elegans* (Wood, WB and the Community of *C. elegans* researchers, Eds.), pp. 337–391, Cold Spring Harbor Laboratory.
17. Hedgecock, E., Culotti, J., and Hall, D. (1990) *Neuron* **2,** 61–85.
18. Peter, M., Neiman, A., Park, H., van Lohuizen, M., and Herskowitz, I. (1996) *EMBO J.* **15,** 7046–7059.
19. Leberer, E., Wu, C., Leeuw, T., Fourest-Lieuvin, A., Segall, J., and Thomas, D. (1997) *EMBO J.* **16,** 83–97.
20. Leeuw, T., Wu, C., Schrag, J., Whiteway, M., Thomas, D., and Leberer, E. (1998) *Nature* **391,** 191–195.
21. Lu, W., Katz, S., Gupta, R., and Mayer, B. (1997) *Curr. Biol.* **7,** 85–94.