Organization and alternative splicing of the *Caenorhabditis elegans* cAMP-dependent protein kinase catalytic-subunit gene (*kin-1*)

Mohammad TABISH*, Roger A. CLEGG†, Huw H. REES* and Michael J. FISHER*¹

*School of Biological Sciences, University of Liverpool, P.O. Box 147, Liverpool, L69 7ZB, U.K., and †Hannah Research Institute, Ayr KA6 5HL, Scotland, U.K.

The cAMP-dependent protein kinase (protein kinase A, PK-A) is multifunctional in nature, with key roles in the control of diverse aspects of eukaryotic cellular activity. In the case of the free-living nematode, *Caenorhabditis elegans*, a gene encoding the PK-A catalytic subunit has been identified and two isoforms of this subunit, arising from a C-terminal alternative-splicing event, have been characterized [Gross, Bagchi, Lu and Rubin (1990) J. Biol. Chem. **265**, 6896–6907]. Here we report the occurrence of N-terminal alternative-splicing events that, in addition to generating a multiplicity of non-myristoylatable isoforms, also generate the myristoylated variant(s) of the catalytic subunit that we have recently characterized [Aspbury, Fisher, Rees and Clegg (1997) Biochem. Biophys. Res. Commun.

238, 523–527]. The gene spans more than 36 kb and is divided into a total of 13 exons. Each of the mature transcripts contains only 7 exons. In addition to the already characterized exon 1, the 5'-untranslated region and first intron actually contain 5 other exons, any one of which may be alternatively spliced on to exon 2 at the 5' end of the pre-mRNA. This N-terminal alternative splicing occurs in combination with either of the already characterized C-terminal alternative exons. Thus, *C. elegans* expresses at least 12 different isoforms of the catalytic subunit of PK-A. The significance of this unprecedented structural diversity in the family of PK-A catalytic subunits is discussed.

Key words: exon, isoform, myristoylation, nematode, transcript.

INTRODUCTION

cAMP-dependent protein kinases (protein kinase As, PK-As) occur widely throughout eukaryotic organisms [1]. A considerable degree of sequence homology exists between the PK-A catalytic subunits found in different organisms; additionally, in multicellular organisms, N-myristoylation of the catalytic subunit is an almost universal covalent modification [2]. It is generally believed that the function of N-myristoylation is to contribute to the generation of a 'targeting signal', pre-disposing the myristoylated protein to associate either with a cellular membrane or other soluble protein(s) [3]. In the case of the PK-A catalytic subunit, there is some evidence for the selective utilization of phosphorylatable protein substrates by the myristoylated catalytic subunit, when compared *in vitro* with non-myristoylated (or non-myristoylatable) enzyme [4,5]. This may imply that myristoylated catalytic subunit is necessary for a subset of the many cellular functions regulated by this enzyme. If this were the case, those multicellular organisms that express non-myristoylatable PK-A catalytic subunits would be anticipated also to express myristoylated variant(s). This has been confirmed for the sea slug, *Aplysia*, which expresses multiple isoforms of PK-A catalytic subunit. These arise from an alternative-splicing event in the coding region of the N-terminus in combination with either of two internal sequences, which result from alternative splicing of mutually exclusive exon cassettes [6]. In *Aplysia*, one N-terminal variant resembles the paradigm N-myristoylatable sequence of the murine α -catalytic subunit. The other N-terminal variant resembles the *Saccharomyces cerevisiae* catalytic subunit homologue encoded by the *TPK1* gene. In the cases of the slime mould *Dictyostelium discoideum* [7] and the free-living nematode *Caenorhabditis elegans* [8] the situation is quite different: distinct non-myristoylated N-termini (lacking a myristoylatable glycine residue at position 1 in the polypeptide) occur in the PK-A catalytic subunits expressed by these organisms. In *C. elegans*,

sequence analysis has revealed that both the described catalytic-subunit isoforms arise from a single splicing event near the C-terminus. Whereas these are homologous over most of their sequence, to the mammalian α -catalytic subunit, the N-terminus is similar to the non-myristoylatable 'N2 polypeptide' seen in *Aplysia* [6]. These observations contrast with our recent immunodetection of an N-myristoylated catalytic-subunit polypeptide in *C. elegans* [9]. Given that there is reported to be only one gene encoding PK-A catalytic subunit in *C. elegans* [8], it is possible that both N-myristoylated and unmyristoylatable variants arise as a consequence of a 5' alternative-splicing event in pre-mRNA. We have, therefore, re-examined the *C. elegans* PK-A catalytic-subunit gene (*kin-1*) in order to provide an explanation for the molecular mechanism by which the N-myristoylated catalytic subunit arises in *C. elegans* and, consequently, provide a more complete understanding of the overall patterns of PK-A catalytic-subunit expression in *C. elegans*.

MATERIALS AND METHODS

Materials

The Expand[™] Reverse Transcriptase, Expand[™] High Fidelity PCR system and Klenow fragment were from Boehringer Corp. (London, U.K.). TOPO TA cloning vector was from Invitrogen BV (Leek, The Netherlands). The Hybaid Recovery Plasmid MiniPrep kit and Hybaid Recovery DNA-purification kit II were obtained from Hybaid Ltd. (Teddington, Middx., U.K.). Quickhyb solution was from Stratagene (La Jolla, CA, U.S.A.). The Dynabead mRNA purification kit was from Dynal (U.K.), Ltd. (Bromborough, Wirral, U.K.). The low-DNA-mass ladder was from Gibco BRL (Paisley, Scotland, U.K.). [α -³²P]dCTP was from Amersham International (Amersham, Bucks., U.K.). Oligonucleotides were synthesized using in-house facilities.

Abbreviations used: PK-A, protein kinase A (cAMP-dependent protein kinase); RT, reverse transcriptase; EST, expressed sequence tag.

¹ To whom correspondence should be addressed (e-mail fishermj@liv.ac.uk).

Animals

The Bristol N2 wild-type strain of *C. elegans*, from the MRC Laboratory of Molecular Biology, Cambridge, U.K. was used in all experiments. Asynchronous cultures of *C. elegans* were grown and maintained on *Escherichia coli* (OP50) essentially as described in [10]. Viability of nematodes was routinely assessed by visual inspection using a light microscope.

Preparation of RNA

Total RNA was isolated from mixed-stage nematodes using a modification of the method described in [11]. Briefly, nematodes were homogenized, using an Ystral disperser, in a 2:1:1 mixture of lysis buffer [4 M guanidine isothiocyanate/0.13 % (v/v) Sarkosyl/33 mM Tris/HCl (pH 8.0)/0.5 % 2-mercaptoethanol/6.7 mM EDTA]/phenol/chloroform. After centrifugation at 10000 g for 20 min at 4 °C the aqueous phase was removed and nucleic acid was precipitated by the addition of 0.025 vols. of 1 M acetic acid/0.75 vols. of ethanol. The precipitated nucleic acid was redissolved in TE buffer (10 mM Tris/HCl, pH 8.0/1 mM EDTA), then RNA was precipitated selectively by addition of an equal vol. of 5 M LiCl/20 mM EDTA (pH 8.0) [12]. The resulting suspension was incubated on ice for 45 min and then centrifuged at 12000 g for 30 min at 4 °C. The RNA precipitate was washed with 2.5 M LiCl/10 mM EDTA (pH 8.0) and then washed with 75 % (v/v) chilled ethanol. Finally, the RNA was redissolved in diethyl pyrocarbonate-treated water.

Reverse transcriptase (RT)-PCR

Transcripts encoding the PK-A catalytic subunit were identified using RT-PCR. *C. elegans* total RNA (3 µg) was primed with oligo(dT)₁₅, and single-stranded cDNA was synthesized using Expand[™] Reverse Transcriptase at 42 °C for 1 h (total vol. 20 µl). Subsequently, 1 µl of the single-stranded-cDNA preparation was added to Expand[™] High Fidelity PCR system buffer for PCR using appropriate sets of primers (total vol. 50 µl). PCR was performed for 30 cycles (denaturation, 94 °C, 3 min, 1 cycle; denaturation, 94 °C, 1 min; annealing, 55 °C, 1 min; extension, 68 °C, 2 min; finally, 68 °C for 7 min). RT-PCR product (5 µl) was subjected to electrophoresis on a 1 % (w/v) agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

Nested RT-PCR

For transcripts of low abundance, for which the RT-PCR products could not be visualized after the procedure described above, nested PCR was performed. After the RT reaction, as described above, PCR amplification was carried out for 30 cycles (as above) using oligo(dT)₁₅ and first-exon-specific primers as down- and upstream primers, respectively. The resulting RT-PCR product (1 µl) was used as a template for further amplification by PCR using the same upstream primer and a downstream primer from either exon 4, exon N or exon 7. The resulting nested PCR product (5 µl) was then subjected to electrophoresis on a 1 % (w/v) agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

Subcloning and sequencing of RT-PCR products

RT-PCR products were fractionated on 1 % (w/v) agarose gels. Appropriate ethidium-bromide-stained bands were cut from the gels and DNA was purified using a Hybaid DNA-purification kit. The purified DNA was subcloned using the TOPO TA cloning vector. *E. coli* JM109-competent cells were transformed.

Transformed colonies were grown overnight at 37 °C. Plasmid DNA was purified using a Hybaid plasmid-purification kit. Plasmids containing the insert were sequenced using an automatic sequencer (ABI Prism 377) by the dideoxy chain-termination procedure [13] using either M13 forward or reverse primers.

Northern-blot analysis

Poly(A)⁺ RNA was prepared from total RNA using a Dynabead mRNA purification kit. Poly(A)⁺ RNA (2.5 µg) was denatured with 2.2 M formaldehyde/50 % (v/v) formamide and was subjected to electrophoresis using 1 % (w/v) denaturing agarose gels. RNA was then transferred to nylon membrane and cross-linked by UV irradiation. Exon-specific probes were prepared using small RT-PCR products (containing the appropriate first exon, as confirmed by sequence analysis) as a template. Each template was primed with a 23mer of exon 2 (close to the first exon-exon 2 junction) and extended by Klenow fragment in the presence of [α -³²P]dCTP, appropriate reaction buffer and other nucleotides, at 37 °C for 30 min. Pre-hybridization was carried out for 20 min at 68 °C in Quickhyb solution. Hybridization was carried out for 1 h at 68 °C. Subsequent washing was performed twice in 2 × SSC (1 × SSC is 0.15 M NaCl/0.015 M sodium citrate)/0.1 % SDS for 20 min at room temperature and then twice in 0.1 × SSC/0.1 % SDS for 30 min at 60 °C. The blots were autoradiographed at −70 °C using Fuji medical X-ray film.

Bioinformatics tools

The databases used for searches were those at the *C. elegans* Sequencing Project Centre (http://www.sanger.ac.uk/Projects/C_elegans/blast_server.shtml), from the *C. elegans* sequencing project, a collaboration between the Sanger Centre, Cambridge, U.K., and the Genome Sequencing Center at the Washington University School of Medicine, St. Louis, MO, U.S.A., and Yuji Kohara's *C. elegans* expressed sequence tag (EST) database (http://www.ddbj.nig.ac.jp/htmls/c-elegans/html/CE_INDEX.html). The Gene Finder tool was at <http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html>. Alignment analysis was carried out using the Genestream Align tool at <http://www2.igh.cnrs.fr/bin/align-guess.cgi>. Primers were either designed manually or using the tool available at <http://www.williamstone.com/primers>.

RESULTS AND DISCUSSION

Alternative splicing of the *C. elegans* PK-A catalytic-subunit gene

In order to understand the molecular mechanisms underlying the generation of the myristoylated variant(s) of the *C. elegans* PK-A catalytic subunit, we re-examined the nature of the gene encoding this polypeptide. In the first instance, we searched the *C. elegans* sequencing project databases for sequences showing significant similarities with mammalian PK-A catalytic-subunit sequences. These searches indicated the presence of two potential PK-A catalytic-subunit genes. The first gene (*kin-1*) was that already characterized from chromosome I by Rubin's group [8] and located within cosmids F21F12 (accession no. Z81511) and ZK909 (accession no. Z82096) [14]. In addition, we also identified a gene (F47F2.1) from the X chromosome, located within cosmid F47F2 (accession no. U40943), which is predicted to encode a 'hypothetical PK-A catalytic-subunit-like polypeptide'. Although the occurrence of other genes encoding PK-A catalytic subunits cannot be ruled out, the completion of the *C. elegans* sequencing project suggests that it is quite possible that all the PK-A isoforms present in *C. elegans* are the products of

Figure 1 The structure and organization of the *C. elegans* PK-A subunit gene (*kin-1*)

The intron/exon organization of the *C. elegans* PK-A catalytic subunit, as described by Gross et al. [8] is illustrated in the top line. Exons 1–7 (coding region only) are indicated as solid vertical bars and introns are indicated as interconnecting lines drawn to the same scale as the exons. The C-terminus splice variants described by Gross et al. [8] involve the use of either exon N or exon 7. The exons predicted by Gene Finder analysis of the 5'-untranslated region preceding exon 1 through to the 5'-end of exon 2 are indicated in the bottom line. Here, predicted exons, N'1–N'6, are again indicated as solid bars and introns are indicated as interconnecting lines drawn to the same scale as the exons. The distance between exons N'1 and 7 is ≈ 36 kb. Exon N'3 is identical to exon 1 as shown in the top line. *ter* represents the termination codon for the predicted gene F21F12.1. The sizes and positions of the predicted PK-A catalytic-subunit exons and introns are indicated in Table 1.

Table 1 Details of the predicted *C. elegans* PK-A catalytic-subunit exons and introns

The sizes and positions of the predicted PK-A catalytic-subunit exons and introns are indicated. In addition, each 5'-exon (upper case)–intron (lower case) boundary is indicated. The 3'-intron–exon 2 boundary is attttcagTCAAAGAA. *ter* represents the termination codon for the predicted gene F21F12.1.

Region	Size in bp (amino acids in brackets)	5'-Exon–intron boundary	Cosmid (position)
<i>ter</i> –N'1 untranslated region	3964		F21F12
exon N'1	136 (45.3)	CAGAGACAgtggg	F21F12 (07815–07950)
intron N'1–N'2	4306		F21F12
exon N'2	22 (7.3)	TGTTcATTGgtatt	F21F12 (12257–12278)
intron N'2–N'3	9905		F21F12
exon N'3 (= exon 1)	70 (23.3)	CGGCGTCGcgtatg	F21F12 (22184–22253)
intron N'3–N'4	859		F21F12
exon N'4	103 (34.3)	TAGAACTGgtaat	F21F12 (23113–23215)
intron N'4–N'5	2054		F21F12
exon N'5	157 (52.3)	TGGAGGATTgtaag	F21F12 (25270–25426)
intron N'5–N'6	2549 + 649 = 3198		F21F12/ZK909
exon N'6	40 (13.3)	TCAAGTTTTgtagg	ZK909 (00650–00689)
intron N'6–2	516		ZK909

these two genes. Further, given that searches of cDNA and EST databases did not provide any evidence for the expression of the X-chromosome gene, we initially re-examined the structure of the chromosome 1 gene, *kin-1*.

The cosmids F21F12 and ZK909, containing the genomic region encoding the N-terminal portion of the PK-A catalytic subunit, were subjected to Gene Finder analysis. This tool predicts genes beginning with the initiator methionine codon and ending with a translation stop codon, together with protein-coding exons and splice sites. In a 25 kb genomic sequence, extending from the predicted termination codon of gene F21F12.1 (immediately upstream of the *kin-1* gene), to the 5'-untranslated region before exon 1 and through to the start of exon 2 of the *kin-1* gene (see Figure 1), Gene Finder analysis predicted the occurrence of five additional exons. These exons were predicted to encode between 7 and 52 amino acids, each having methionine as the initiation codon and each potentially able to splice into exon 2 (Figure 1 and Table 1).

RT-PCR amplification was used to investigate the possible occurrence of these exons in transcripts. Total RNA, prepared from mixed-stage *C. elegans*, was reverse transcribed and the resulting single-stranded cDNA was PCR amplified. Upstream primers corresponded to each of the potential exons predicted by

Table 2 Primer design for RT-PCR analysis for the occurrence of predicted PK-A catalytic-subunit exons

Exons are named as indicated in Figure 1. Primers were designed as indicated in the Materials and methods section.

Cosmid	Exon	Cosmid position	Size (bp)	Direction
F21F12	N'1	07811–07840	30	Forward
F21F12	N'2	12245–12271	27	Forward
F21F12	N'3 (= 1)	22087–22107	21	Forward
F21F12	N'4	23109–23127	19	Forward
F21F12	N'5	25267–25296	30	Forward
ZK909	N'6	00643–00672	30	Forward
ZK909	4	09499–09521	23	Reverse
ZK909	N	11299–11321	23	Reverse
ZK909	7	11800–11822	23	Reverse

Gene Finder. Downstream primers were either from exon 7 or exon N at the 3' end of the gene or, alternatively, from exon 4 within the internal region of the gene (see Table 2).

Using these primers, we were able to identify RT-PCR products

Lane	PCR Product	Primers set (5', 3')		Predicted size (bp)	Isoform
1	Exon-N'1---Exon-4	N'1	4	369	internal
2	Exon-N'1---Exon-N	N'1	N	1116	1
3	Exon-N'1---Exon-7	N'1	7	1101	2
4	Exon-N'2---Exon-4	N'2	4	272	internal
5	Exon-N'2---Exon-N	N'2	N	1019	3*
6	Exon-N'2---Exon-7	N'2	7	1004	4*
7	Exon-N'3---Exon-4	N'3	4	395	internal
8	Exon-N'3---Exon-N	N'3	N	1142	5 †
9	Exon-N'3---Exon-7	N'3	7	1127	6 †
10	Exon-N'4---Exon-4	N'4	4	335	internal
11	Exon-N'4---Exon-N	N'4	N	1082	7
12	Exon-N'4---Exon-7	N'4	7	1067	8
13	Exon-N'5---Exon-4	N'5	4	388	internal
14	Exon-N'5---Exon-N	N'5	N	1135	9*
15	Exon-N'5---Exon-7	N'5	7	1120	10
16	Exon-N'6---Exon-4	N'6	4	273	internal
17	Exon-N'6---Exon-N	N'6	N	1020	11*
18	Exon-N'6---Exon-7	N'6	7	1005	12

Figure 2 RT-PCR analysis of predicted PK-A catalytic-subunit exons

RT-PCR amplification was used to determine the presence of transcripts, containing predicted exons, in total RNA prepared from mixed-stage *C. elegans* as described in the Materials and methods section. Top panel: the migration of a series of size markers is indicated on the right. Bottom panel: Table showing the nature of each RT-PCR product (lanes 1–18). †, Isoforms reported previously [8]; *, products observed only after nested PCR analysis.

of the sizes anticipated, on the basis of the presence of the predicted exons in transcripts (Figure 2). Furthermore, products corresponding to each of the predicted N-terminal exons in combination with either of the previously reported C-terminal exons (i.e. exon 7 or exon N) were observed. However, it should be noted that certain transcripts were of extremely low abundance and only detectable using nested RT-PCR protocols. In particular, transcripts containing exon N'2 and either of the C-terminal exons (N or 7) and transcripts containing exons N'5 or N'6 in combination with the C-terminal exon N were of low

abundance. It is also apparent from Figure 2 that, in addition to the anticipated PCR products, other products were also observed in the case of the N'4-N and N'4-7 combinations of primers (Figure 2; lanes 11 and 12). When these additional products were purified and sequenced it was apparent that they arose non-specifically, as a consequence of the N'4 primer showing similarity with another distinct region of the genome, unrelated to the PK-A catalytic-subunit gene (results not shown).

Following the identification of RT-PCR products indicating the occurrence of the predicted alternative N-terminal exons

Figure 3 Northern-blot analysis of predicted PK-A catalytic-subunit transcripts

Northern-blot analysis of poly(A)⁺ RNA prepared from mixed-stage *C. elegans*, using probes (based upon the first exon–exon 2 junction) specific for the predicted transcripts for the PK-A catalytic-subunit products, as described in the Materials and methods section. Lane 1, N'1-; lane 2, N'2-; lane 3, N'3-; lane 4, N'4-; lane 5, N'5-; lane 6, N'6-specific probes were used. The migration of a series of size markers is indicated on the left.

spliced to exon 2, we searched the *C. elegans* databases for evidence to support the occurrence of possible N-terminal splicing events. A search of Yuji Kohara's *C. elegans* EST database indicated that cDNA yk252f9 arises from the splicing of exon N'4 with exon 2.

We carried out Northern-blot analysis on poly(A)⁺ RNA to confirm directly the occurrence of mature transcripts containing the predicted exons alternatively spliced directly to exon 2 (Figure 3). This analysis confirmed the presence of each of the anticipated

transcripts (≈ 2.5 kb). In this context, it is of interest that other workers have noted previously the occurrence of multiple mRNA species for the *C. elegans* PK-A catalytic subunit [8], although these were ascribed to the possible use of alternative polyadenylation sites and/or transcription-initiation sites.

To further confirm the existence of the predicted exons within the PK-A catalytic-subunit transcripts, products obtained by RT-PCR, using primers corresponding to the predicted exons at the 5' end and a primer from exon 4 at the 3' end, were subcloned into a TOPO TA cloning vector and sequenced. Sequencing confirmed the existence of each predicted exon spliced into exon 2 and Figure 4 indicates the deduced amino acid sequences for the N-terminus of each isoform. In each case the deduced amino acid sequence was identical to that which could be predicted by Gene Finder and translational analysis.

Other features of the *C. elegans* PK-A catalytic-subunit gene

Following our studies of the 5' terminus of the *C. elegans* PK-A gene, we extended our analysis beyond exon 2. These latter studies did not indicate any alternative splicing or structural features distinct from those reported previously [8]. However, these studies did highlight an unusual feature of this region of the *C. elegans* genome. Using Gene Finder, it was possible to predict the occurrence of two distinct coding regions within the sequence between exon 2 and exon 3 of the PK-A catalytic-subunit gene. First, a gene product (position 7344–6355 on cosmid ZK909), encoded by the strand opposite from that encoding the PK-A catalytic subunit, can be predicted to contain two introns (intron 1 [571 bp] at 7200–7057 and intron 2 [103 bp] at 6627–6522). This gene product has already been predicted (Wormpep entry ZK909.1; from the *C. elegans* protein database WormPep). However, we were able to also predict the presence of a potential coding region, encoded by the same strand as that encoding the PK-A catalytic subunit (position 6406–7287 on cosmid ZK909). This coding region does not appear to be interrupted by introns and, potentially, gives rise to a 293 amino acid polypeptide. A search of EST databases did provide support for the expression of this coding region. One cDNA clone (yk428h7.5) corresponding to the predicted coding region was identified. The sequence of the 5' terminus of this cDNA matched that from the region between the putative coding region and exon 2 of the PK-A catalytic-subunit gene. Additionally, this cDNA contained the sequence of the 103 bp intron predicted for a gene encoded by the opposite strand (see above). Taken together, these features suggest that transcripts may arise from within the region between

Exon	Deduced amino acid sequence	Acc. No.
N'1	<u>MGNAASGGSSGGGSARRGNGGGNNGSDYNNAMVFSNGRLAAAE</u> T <i>KEFLDKAREDFKQRW</i>	AJ011936
N'2	MGSMVFI V <i>KEFLDKAREDFKQRW</i>	AJ011937
N'3 (1)	MLKFLKPKSSDEGSSKDNKNSAS L <i>KEFLDKAREDFKQRW</i>	AJ011938
N'4	MLSSSFVRGSMKERKNEALKNHKS ^Y ISGGYLET V <i>KEFLDKAREDFKQRW</i>	AJ012354
N'5	MPTRLDIVGNLQFSSSTDNGDEDDQEADV T ACFVLPSPSSFSKLSILDDPVED F <i>KEFLDKAREDFKQRW</i>	AJ012355
N'6	MSSSSNKKVQVK F <i>KEFLDKAREDFKQRW</i>	AJ012356

Figure 4 Amino acid sequences encoded by alternatively spliced N-terminal exons

Deduced amino acid sequences encoded by alternative first exons (shown in regular font) and the first part of the amino acid sequence encoded by exon 2 (shown in italics). Amino acids encoded by a codon at the splice site are shown in bold. The potential myristoylation site is underlined. The EMBL data library accession numbers are indicated.

(A) RT-PCR amplification was used to determine the presence of transcripts, encoding the F47F2.1 PK-A catalytic-subunit-like product, in total RNA prepared from mixed-stage *C. elegans* as described in the Materials and methods section. The PCR primers were designed manually from cosmid F47F2 to detect a full-length cDNA, including translational start and stop codons for the polypeptide F47F2.1. The migration of a series of size markers (lane 1) is indicated on the left. (B) Northern-blot analysis of poly(A)⁺ RNA prepared from mixed-stage *C. elegans*, using a probe specific for transcripts encoding the F47F2.1 PK-A catalytic-subunit-like product, as described in the Materials and methods section. The migration of a series of size markers is indicated on the left. (C) Alignment of the deduced amino acid sequence of the F47F2.1 polypeptide (CE04603) with the murine α -catalytic subunit. Alignment analysis was carried out using the Genestream Align tool. Two dots represent identity, one dot represents conservation and no dots indicates non-identity. Residues shown in bold are the 12 key conserved residues [18]. The EMBL data library accession numbers are indicated at the 3' ends of the sequences.

exons 2 and 3 of the PK-A catalytic-subunit gene (position 6406–7287 on cosmid ZK909), from the same strand as the PK-A catalytic-subunit gene. It is possible that the presence of this sequence within the EST database reflects the occurrence of contaminating PK-A catalytic-subunit pre mRNA. However, the relatively small size of the yk428h7.5 cDNA clone suggests that it is not derived from a pre mRNA.

The nature of this potential gene product is unclear. Searches of the databases, using the predicted polypeptide sequence, did not reveal significant similarities to any other polypeptides. Although unusual, the occurrence of a gene within an intronic region of another gene has been noted before in *C. elegans*. For example, the *spe-26* gene involved in spermatid formation appears to be contained within an intron of a somatically expressed gene that is transcribed in the opposite direction from *spe-26* [15].

Expression of the PK-A catalytic-subunit-like gene (F47F2.1)

Although we have focused our attention on the *kin-1* gene, we also carried out limited analysis of the putative F47F2.1 gene product. As described above, there was no evidence from searches of the EST databases for the expression of this gene. Nonetheless, using primers designed to uniquely recognize the predicted mRNA from this gene, we carried out RT-PCR analysis. This revealed an extremely abundant product of the anticipated size, indicative of expression of this gene (Figure 5A; lane 2). These observations were substantiated by the sequencing of this product and by Northern-blot analysis of poly(A)⁺ RNA (Figure 5B). Taken together, these observations suggest that this gene does express a transcript that may encode a PK-A catalytic-subunit-like polypeptide. There was no evidence for any form of alternative splicing of this gene. Database analysis revealed that the F47F2.1 gene product shared significant homology ($\approx 55\%$ identity) with a PK-A catalytic subunit from the pig roundworm (*Ascaris suum*) [16]. Figure 5(C) shows the alignment between the amino acid sequence of the F47F2.1 gene product and the murine PK-A α -catalytic subunit (murine C α). This reveals 45% identity, including all of the 12 key conserved residues (shown in bold in Figure 5C) identified within the eukaryotic protein-kinase superfamily [17]. Additionally, the motif YRDLKPEN is identical between the catalytic loop (subdomain VIB) of the murine C α subunit and the F47F2.1 polypeptide (deduced residues 183–190), as are all three glycines in the glycine-rich loop region. Also in the vicinity of the catalytic loop is extensive similarity/conservation between subdomains VII, VIII and IX of murine C α and deduced residues 194–250 of the F47F2.1 polypeptide. In particular, the threonine residue of the activation loop (Thr-197 in murine C α) and its neighbouring residues are conserved in the *C. elegans* polypeptide. Phosphorylation of Thr-197 in the PK-A catalytic subunit (and its homologues in other kinases) is required for the expression of catalytic activity; by analogy, we anticipate that Thr-216 of the *C. elegans* deduced sequence will be found to be phosphorylated in the mature expressed protein. In contrast, the other major constitutive phosphorylation site (Ser-338) in murine C α is not conserved in the *C. elegans* sequence. Outside the catalytic core region, the amphipathic helix comprising residues 15–35 in murine C α has no obvious analogue in the *C. elegans* sequence. This is of interest in that it appears to violate the principles of the general model proposed for eukaryotic protein kinases by Veron et al. [18]. This model predicts that an A-helix motif, complementary to the surface of the conserved catalytic core, should occur in the kinase sequence; important functional roles in the maintenance of solubility, stability and domain interactions have been attributed to this A-helix in the mammalian catalytic subunit

[19]. However, the F47F2.1 polypeptide has no such A-helix in its N-terminal sequence, either located contiguously or, as in the case of the *D. discoideum* catalytic subunit [7], non-contiguously to the core. Neither can any candidate A-helix be found in a non-contiguous C-terminal location, relative to the core of the F47F2.1 polypeptide (see Figure 5C). Thus for this polypeptide, conformity with the complementary A-helix model can be secured only by hypothesizing that it comprises the catalytic subunit of a multi-subunit kinase and that the complementary A-helix exists on a separate subunit. Recent evidence obtained from studies of the *Dictyostelium* catalytic subunit may indicate, however, that the complementary A-helix is not an invariant necessity for protein-kinase function; mutants lacking it were fully biologically active in this organism [7]. Activity was abolished, however, in this enzyme, by deletion of a 27-residue hydrophobic sequence flanking the N-terminal boundary of the core [20]. Of the 20 amino acids immediately adjacent to the core, 13 were apolar, leading these authors to postulate that such a flanking sequence could serve the same function as that proposed for the A-helix. Expression studies will be needed to examine the implications for kinase function of the absence of such a markedly hydrophobic sequence (seven apolar residues out of 20) N-terminal to the core boundary in the F47F2.1 polypeptide.

A juxta-C-terminal cluster of acidic residues (Asp-328–Glu-334) in murine C α constitutes the recognition/cleavage site for a highly specific membrane-associated protease [21] that cleaves and inactivates free native catalytic subunit and has been proposed to play a role in the regulation of kinase activity. A conserved region spanning positions 340–352 in the *C. elegans* sequence includes six of these seven residues. This conserved region also occurs in all those members of the *kin-1* splice-variant family described here in which exon 7 but not exon N is transcribed.

The significance of the occurrence of multiple forms of PK-A catalytic subunit in *C. elegans*

In summary, we have provided evidence for the occurrence of transcripts encoding multiple isoforms of the PK-A catalytic subunit in *C. elegans*. As a consequence of N-terminal splicing, in combination with the known C-terminal splicing event, there are as many as 12 different isoforms derived from the *kin-1* gene. Of the transcripts for these different isoforms, eight appear to be relatively abundant whereas four are of considerably lower abundance (see Figure 2). In addition, transcripts arising from the PK-A catalytic-subunit-like gene (F47F2.1) also appear to be abundant.

Analysis of the deduced amino acid sequences for the different isoforms is informative (Figure 4). Isoforms involving exon N'1 are likely to be targets for N-terminal myristoylation, as they contain a classical myristoylation motif. It is therefore entirely possible to explain the occurrence of myristoylated [9] and unmyristoylated catalytic subunits in *C. elegans* on this basis. Searches of the databases indicated that there was a significant degree of homology between the deduced amino acid sequence of exon N'1 and a PK-A catalytic subunit (accession no. 125215) from *Drosophila* [22]. Transcripts encoding exon N'2 appear to be of low abundance (see Figure 3). This exon is the smallest one encoded (7 amino acids) by this gene. Although this exon contains a glycine residue adjacent to the N-terminal methionine, it would appear to be an unlikely target for myristoylation, lacking an obvious consensus sequence. Transcripts encoding exon N'3 have been characterized previously [8]; these are clearly not targets for myristoylation and appear to share features with the unmyristoylatable catalytic subunits found in *Aplysia* and the

S. cerevisiae TPK1 gene (see the Introduction). In addition, a dog hookworm (*Ancylostoma caninum*) PK-A catalytic subunit shows a high degree of homology with N'3-containing isoforms [23]. Transcripts encoding N'4 are also clearly not targets for myristoylation. Furthermore, transcripts encoding this exon are abundant enough to have been picked up in EST libraries (see above). In contrast, transcripts involving exons N'5 and N'6 are less abundant; specifically, those transcripts encoding either of these exons in combination with the C-terminal exon N are of low abundance. The marked N-terminal plasticity of the *kin-1* gene products that we have described does not involve the A-helix motif (discussed above). This is encoded by exon 2 and is intact in all 12 putative *kin-1* isoforms. Structural analysis of the murine $C\alpha$ isoform by X-ray crystallography has indicated that the N-terminal sequence comprising amino acid residues 1–14, encoded by exon 1, adopts a disordered configuration when not anchored on to the kinase surface by the addition of an N-myristoyl substituent [24]. Similarly, an epitope within this sequence is more accessible to its cognate antibody in recombinant non-myristoylated catalytic subunit than in the normally N-myristoylated protein [25]. These observations suggest possible functions for non-myristoylated variant N-termini in mediating inter-molecular interactions between catalytic subunit and partner molecules, perhaps leading to targeting or tethering of catalytic subunit. An example of this is afforded by the recent report linking the atypical subcellular distribution of catalytic subunit in ovine spermatozoa to the expression, in this cell type, of a variant protein (C_v) with a non-myristoylatable N-terminus. It is likely that this variant arises by alternative splicing of a novel exon 1 in the $C\alpha$ gene [26]. In the case of the *C. elegans kin-1* isoforms, it is noteworthy that the sequences deduced from exons N'2–N'6 are markedly different from one another in, for instance, their charge properties and might therefore mediate interactions with different populations of partner molecules. One known interaction partner that deserves consideration in this context is the regulatory (R-) subunit, which is an RI type [1] in *C. elegans*. This again raises, indirectly, the issue of targeting of PK-A molecules, since R-subunit anchoring via AKAPs (A-kinase-anchoring proteins) [27] in this organism has recently become a possibility following the cloning of a novel RI-selective AKAPCE [28].

The PK-A catalytic-subunit structural diversity indicated by the present studies is unprecedented. To what extent this reflects functional specialization, rather than redundant diversity, remains to be established. As indicated earlier (see the Introduction and Discussion above), the expression of both myristoylated and non-myristoylated (non-myristoylatable) variants may be of considerable significance in the differential phosphorylation of distinct sub-sets of target polypeptides and, therefore, may be of key importance in the control of specific cellular activities. Likewise, the different non-myristoylatable variants may also show significant differences in their temporal and/or spatial expression in *C. elegans*. This may also lead to differential

phosphorylation of distinct sub-sets of target polypeptides and, consequently, the control of specific cellular activities.

M. T. is grateful to the Leverhulme Trust for the provision of a Visiting Fellowship. This work was supported in part by the Scottish Office Agriculture, Environment and Fisheries Department.

REFERENCES

- Hardie, G. and Hanks, S. (eds.) (1995) *The Protein Kinase Facts Book: Protein Serine Kinases*, Academic Press, London
- Carr, S. A., Biemann, K., Shoji, S., Parmelee, D. C. and Titani, I. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6128–6131
- Boutin, J. A. (1997) *Cell. Signalling* **9**, 15–35
- Panchal, R. G., Cheley, S. and Bayley, H. (1994) *J. Biol. Chem.* **269**, 23722–23730
- Adie, E. J., Thomas, P. H., Munday, M. R. and Clegg, R. A. (1995) *Biochem. Soc. Trans.* **23**, 451S
- Beushausen, S., Lee, E., Walker, B. and Bayley, H. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1641–1645
- Etchebehere, L. C., van Bemmelen, M., Anjard, C., Traincard, H., Assemat, K., Reymond, C. and Véron, M. (1997) *Eur. J. Biochem.* **248**, 820–826
- Gross, R. E., Bagchi, S., Lu, X. and Rubin, C. S. (1990) *J. Biol. Chem.* **265**, 6896–6907
- Aspbury, R. A., Fisher, M. J., Rees, H. H. and Clegg, R. A. (1997) *Biochem. Biophys. Res. Commun.* **238**, 523–527
- Brenner, S. (1974) *Genetics* **77**, 71–94
- Conrad, R., Thomas, J., Spieth, J. and Blumenthal, T. (1991) *Mol. Cell. Biol.* **11**, 1921–1926
- Cathala, G., Savouret, J. F., Mendez, B., West, B. L., Karin, M., Martial, J. A. and Baxter, J. A. (1983) *DNA* **2**, 329–335
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J. et al. (1994) *Nature (London)* **368**, 32–38
- Varkey, J. P., Muhlrad, P. J., Minniti, A. N., Do, B. and Ward, S. (1995) *Genes Dev.* **9**, 1074–1086
- Jung, S., Hoffmann, R., Rodriguez, P. H., Mutzel, R. and Hofer, H. W. (1995) *Eur. J. Biochem.* **232**, 111–117
- Hanks, S. K. and Hunter, T. (1995) in *The Protein Kinase Facts Book: Protein Serine Kinases* (Hardie, G. and Hanks, S., eds.), pp. 7–47, Academic Press, London
- Veron, M., Radzio-Andzelm, E., Tsigelny, I., Ten Eyk, L. F. and Taylor, S. S. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10618–10622
- Herberg, F. W., Zimmermann, B., McGlone, M. and Taylor, S. S. (1997) *Protein Sci.* **6**, 569–579
- Dammann, H., Traincard, F., Anjard, C., van Bemmelen, M. X. P., Reymond, C. and Véron, M. (1998) *Mech. Dev.* **72**, 149–157
- Chestukhin, A., Litovchick, L., Muradov, K., Batkin, M. and Shaltiel, S. (1997) *J. Biol. Chem.* **272**, 3153–3160
- Foster, J. L., Higgins, G. C. and Jackson, F. R. (1988) *J. Biol. Chem.* **263**, 1676–1681
- Hawdon, J. M., Jones, B. F. and Hotez, P. J. (1995) *Mol. Biochem. Parasitol.* **69**, 127–130
- Zheng, J., Knighton, D. R., Xuong, N.-H., Taylor, S. S., Sowadski, J. M. and TenEyck, L. F. (1993) *Protein Sci.* **2**, 1559–1573
- Beattie, J. and Clegg, R. A. (1993) *Biochem. Soc. Trans.* **21**, 398S
- San Agustín, J. T., Leszyk, J. D., Nuwaysir, L. M. and Witman, G. B. (1998) *J. Biol. Chem.* **273**, 24874–24883
- Faux, M. and Scott, J. D. (1996) *Cell* **85**, 9–12
- Angelo, R. and Rubin, C. A. (1998) *J. Biol. Chem.* **273**, 14633–14643

Received 5 November 1998/7 January 1999; accepted 27 January 1999