Cloning, Characterization, and Expression of the Gene for the Catalytic Subunit of cAMP-dependent Protein Kinase in Caenorhabditis elegans

IDENTIFICATION OF HIGHLY CONSERVED AND UNIQUE ISOFORMS GENERATED BY ALTERNATIVE SPlicing*

(Received for publication, October 6, 1989)

Robert E. Gross†, Srilata Bagchi, Xiangyi Lu, and Charles S. Rubin
From the Department of Molecular Pharmacology, Atran Laboratories, Albert Einstein College of Medicine, Bronx, New York 10461

The nematode Caenorhabditis elegans expresses substantial amounts of several forms (Mₐ values = 39,000–41,000) of the catalytic subunit (C) of cAMP-dependent protein kinase. Approximately 65% of the total cAMP-dependent phosphotransferase activity is recovered in particulate fractions of homogenates prepared from asynchronous populations of C. elegans. The C subunit is expressed at a low level in cytosolic and particulate compartments during embryogenesis. As the nematodes progress from late embryonic stages to the newly hatched, first larval (L1) stage, C subunit content increases 15-fold. High levels of C subunits are observed in several subsequent larval and adult stages of development. Since the relative abundance of C subunit mRNA changes little with development, it appears that control of C expression is exerted at the translational and/or post-translational levels.

cDNAs for two types of C have been cloned and sequenced. The derived amino acid sequence of a major isoform (CeCATα, 358 residues) is highly homologous (82% identical) with the murine Ca subunit. A second, novel C subunit (CeCATα', 374 residues) has a unique 56-residue carboxyl-terminal region that is generated by the alternative splicing of the C pre-mRNA. The splicing process yields CeCATα' is unusual because it converts the central portion of an apparent l-kilobase (kb) intron to an exon. The alternative exon introduces the novel carboxyl terminus and a new translation stop signal, while simultaneously converting the coding sequence for 40 carboxyl-terminal residues in CeCATα into 3' untranslated nucleotides. The 5' end of the C. elegans C subunit mRNA is produced by the trans-splicing of the C gene transcript to a 22-base pair C. elegans leader sequence originally described by Krause, M., and Hirsh, D. (1987) Cell 49, 753–761.

The 20-kb C. elegans C gene is divided into seven exons by introns ranging in size from 54 to 8000 bp. The sizes of the C. elegans C subunit gene, cytoplasmic mRNA (2.5 kb), and subunit protein are similar to the sizes of the murine Ca gene, mRNA, and polypeptide. However, the nematode and murine C genes differ significantly in the organization of their introns and exons.

Protein kinases regulate the activity of certain enzymes and the functions of other proteins by catalyzing the phosphorylation of serine, threonine, or tyrosine residues in target substrate sequences (1, 2). Upon activation by intracellular second messengers or by extracellular effectors, protein kinases amplify, integrate, and coordinate cellular responses to signals from the external and internal environments of cells. Protein kinases are also central components of hormone, growth factor, and neurotransmitter-activated signal transduction pathways by which cells modulate their metabolism, growth rate, and gene expression during the fundamental processes of differentiation and development. In order to understand how protein kinases participate in differentiation and development it is advantageous to study the structure, function, and regulation of phosphotransferases in organisms of varying complexity that are amenable to biochemical, genetic, and physiological manipulations.

The nematode Caenorhabditis elegans (C. elegans) provides a system with many advantages for the study of the molecular events underlying differentiation and development. The organism is composed of a relatively small number of cells (~1000 cells/nematode), but the cells develop along several pathways to generate highly differentiated digestive, nervous, muscular, hypodermal, and reproductive systems (3–5). The complete lineage of every cell in the adult nematode is known and is precisely reproduced in each animal. Molecular genetic techniques are well advanced (5), and the physical map of the C. elegans genome may soon be completed (6). Genomic screening and characterization are facilitated by a small genome size (10⁸ bp)¹ and the occurrence of small introns in certain genes (7–9). Techniques for construction of transgenic nematodes are being developed and becoming effective (5, 10).

¹ The abbreviations used are: bp, base pair(s); kb, kilobase pair; C, catalytic subunit of cAMP-dependent protein kinase; R, regulatory subunit of cAMP-dependent protein kinase; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl, 15 mM sodium citrate, pH 7.0; SL1, trans-spliced leader sequence at the 5' end of some C. elegans mRNAs; CeCATα, the isoform of C. elegans C subunit that is highly homologous with mammalian C subunits; CeCATα', the isoform of C. elegans C subunit that contains a novel carboxyl terminus as the result of alternative splicing; PCR, polymerase chain reaction.
Moreover, recent studies suggest that some cell-cell interactions central to C. elegans development are mediated by signal transduction systems similar to those in mammals (6, 11–13).

Since there are no reports of investigations on Ser/Thr protein kinases in C. elegans, our first goal in studying phosphotransferases in development was to determine the properties and regulation of a protein kinase that is likely to be essential for normal C. elegans ontogeny. C. elegans neurons contain dopamine, serotonin, and octopamine, and the external application of these agents elicits potent physiological effects in the nematodes (14, 15). Since complexes of these agents with their specific receptors activate adenylate cyclase which contains catalytic (C) and regulatory (R)* subunits of CAMP-dependent protein kinase in C. elegans.

Pertinent questions to be addressed in initial studies on C are: what are the number and nature of C. elegans C polypeptides and their cognate mRNAs? Is the expression of C developmentally regulated? If so, at what level? How is the C gene(s) organized and controlled?

In addition, C. elegans represents a branch of evolution and level of cellular organization that are different from the systems in which CAMP-dependent protein kinases have been intensively studied (e.g. mammals, aplysia, and yeast). Thus, new information on essential conserved regions of C and domains of divergence might also be obtained in these studies.

In this paper we demonstrate that C. elegans expresses substantial amounts of several forms of C subunits. The amount of C subunits increases 15- to 25-fold during the transition from the embryo to the first larval (L1) stage. Control is exerted at the translational and/or posttranslational levels.

cDNAs for two types of C have been cloned and sequenced. The derived amino acid sequences reveal one form of C that is highly homologous with mammalian C subunits and a second, novel C with a unique carboxyl-terminal region produced by alternative splicing of the C pre-mRNA. The 5' end of the C subunit mRNA is generated by trans-splicing of the C gene transcript to a 22-bp leader sequence.

**EXPERIMENTAL PROCEDURES**

**RESULTS**

*Cloning and Characterization of cDNAs for the C Subunit of C. elegans CAMP-dependent Protein Kinase—*A cDNA insert which encodes a portion of the mouse Ca catalytic subunit (16) was radiolabeled by random priming (25) and hybridized with Southern blots of C. elegans genomic DNA at low stringency. Hybridization and washing conditions (see “Experimental Procedures”) were adjusted to obtain an optimal signal-to-background ratio (data not shown). Then these conditions were used to screen a C. elegans λgt10 cDNA library, provided by Dr. Barbara Meyer, MIT, Cambridge, MA. A single 1.7-kb clone was isolated from 500,000 plaques and designated XCl. Fig. 1 shows a Southern blot of C. elegans DNA fragments hybridized with a 32P-labeled, 1.4-kb EcoRI fragment of XCl cDNA under high stringency conditions. The C. elegans and mouse (not shown) probes hybridize with the same restriction fragments. Furthermore, there appears to be only one C. elegans gene that is highly homologous to the C. elegans cDNA probe (Fig. 1).

**Phosphotransferase Activity—**Cytosol and particulate fractions of C. elegans were assayed for CAMP-stimulated phosphotransferase activity using histone H2B as a substrate.

**Approximately 65% (64 ± 7%, n = 5) of the total phosphotransferase activity was recovered in the particulate fraction. The association of a significant level of C. elegans cAMP-dependent protein kinase with the particulate fraction parallels the atypical situation in mammalian brain (40).**

**The Primary Structure of the C. elegans C Subunit—**The cDNA sequence and derived amino acid sequence of C. elegans C are presented in Fig. 3. A comparison of the deduced primary sequences of C. elegans C and mouse Cα is shown in Fig. 4. The C. elegans subunit contains 358 amino acid residues (excluding the initiator Met) compared with 350 amino acids for the mouse protein. Aside from the 8-amino acid extension on the amino terminus of C. elegans C, the two polypeptides align in a colinear manner (Fig. 4). Overall, there is 82% identity of the primary sequences, indicative of a high degree of conservation of the C subunit between nematodes and mammals. All of the hallmarks of protein kinases are preserved in the C. elegans protein, including the ATP-binding site at amino acid residues 59–63, the DLKPEN sequence (residues 174–179) that is diagnostic for Ser/Thr protein kinases (41). the DFG sequence (residues 192–194) that is thought to participate in ATP binding, and the APE catalytic domain indicator sequence at residues 215–217. Furthermore, three other structural motifs of the prototypic catalytic subunit (41) are preserved in the C. elegans protein. These include the catalytic site lysine (40), a potential phosphothreonine at Thr203, corresponding to Thr397 and Thr391 of the mouse and yeast C sequences, respectively. This Thr residue is essential for the interaction of C with the regulatory subunit in yeast CAMP-dependent protein kinase (42). A potential phosphoserine site corresponding to Ser338 of the mouse sequence is found at Ser345 in C. elegans C.
Fig. 3. The cDNA and derived amino acid sequences of the C. elegans C subunit. The numbering of the nucleotide sequence is presented on the right, whereas amino acid residue positions are given above the corresponding residues. Components of the ATP-binding site are enclosed in rectangles; a highly conserved element of the catalytic site is overlined; and a sequence conserved in Ser/Thr protein kinases is underlined.

Cloning and Characterization of the C. elegans C Gene—Overlapping clones of the catalytic subunit gene were isolated from a λEMBL4 gene library by hybridization with a 32P-labeled C subunit cDNA insert (Fig. 4). MG1 and MG2 span the entire coding region, and their alignment is shown in Fig. 2B. An additional clone, pB0195-1, extends further 5' and was derived from cosmid B0195 (obtained from Dr. A. Cousin, Medical Research Laboratory of Molecular Biology, Cambridge, Great Britain). The cDNAs were used to characterize and order the structural gene segments by restriction enzyme fragmentation and Southern analysis. The C gene exons (Fig. 5A) and the smaller introns and all of the intron/exon boundaries were sequenced (Fig. 5A, Table II). The C. elegans catalytic subunit gene encompasses approximately 20 kb of DNA and comprises seven exons interrupted by six introns ranging in size from 54 to 8000 bp (Fig. 5A, Table II). The distribution of large and small introns and exons is highly asymmetric, such that the amino-terminal 43 amino acids are encoded by sequences dispersed within 16 kb of genomic DNA, whereas the next 315 amino acids are encoded by exons within 2.5 kb of DNA (Table II) The splice donor and acceptor sites are shown in Table II; the derived consensus sequences match the accepted C. elegans consensus sequences (43).

The 5' End of the C. elegans C Subunit Transcript Is Derived by Trans-splicing to a Ubiquitous 22-bp Leader Sequence—The 5' end of the untranslated region of the C cDNA is shown in Fig. 6. When sequencing the corresponding region of the gene, we noted that the first 7 base pairs of the cDNA diverged from the gene sequence immediately upstream from a consensus 3' splice acceptor sequence within the gene. It appears that this mismatch was not due to an intron interrupting the 5'-untranslated region because the divergent nucleotides in the cDNA perfectly match the 3' end of one of the C. elegans trans-spliced leader sequences (44). These leaders (e.g. SL1, Fig. 6) are attached to 10% of C. elegans transcripts (45) by a reaction utilizing a 5' splice donor site in the leader transcript and a 3' splice acceptor site in the target transcript. To test whether the mismatched 7 base pairs are derived from the 22-bp leader sequence, we sequenced the 5' end of the C. elegans C subunit mRNA. For the S1 nuclelease experiment, we determined whether C subunit mRNA could protect a 60-bp synthetic antisense oligonucleotide (complementary to CR53, Fig. 6) whose sequence corresponds to the 22-bp leader appended to the complementary oligonucleotide as a positive control to demonstrate that the S1 nuclease was active and efficient. If the C subunit mRNA is modified by trans-splicing to the 22-bp leader, then it would protect 54 bp of the oligonucleotide; otherwise, a much smaller protected fragment would be observed. The 5'-end labeled oligonucleotide was hybridized with C. elegans RNA, digested with S1 nuclease, and the protected probe was resolved on a 6% polyacrylamide gel containing 7 M urea (Fig. 7, lane 1). Three major digestion products of 56, 55, and 54 base pairs were observed. It is likely that the two larger products result from inefficient S1 nuclease digestion on 1- or 2-base pair overlaps. The generation of the 54-56 bp protected fragment and the absence of a 32-bp fragment indicate that a complete 22-bp leader sequence is trans-spliced to the C subunit mRNA.
other products might be derived from primary transcripts of the C gene or C mRNA modified in another way or reflect a premature termination site in the reverse transcriptase reaction.

**Alternative Splicing of the C. elegans Gene Transcript Generates a Novel C mRNA**—As mentioned above, three cDNAs for the C subunit displayed restriction fragment size heterogeneities at their 3' ends. These regions were mapped and sequenced, and their structures are diagrammed in Fig. 2A. In areas of overlap, XCl has the same structure as XCl (Fig. 2A), except for the preservation within the XCl cDNA of intron D (Fig. 5). Since the presence of this intron interrupts the open-reading frame of the cDNA after amino acid 119, we believe XCl to be derived from an incompletely spliced precursor of the C subunit mRNA.

The structure of XCl is unique and interesting. Restriction mapping revealed that XCl contained an insertion of approximately 250 bp towards the 3' end. Upon sequencing we discovered that the insertion was derived from an internal region of the 1-kb intron F (Fig. 5). Since the 5' and 3' ends of intron F were not found within the cDNA, XCl could not be a copy of an incompletely spliced nuclear precursor as this would generate a cDNA containing the entire intron. Furthermore, the inserted sequence is surrounded by consensus splice donor and acceptor sites and preserves an open-reading frame that continues for 57 amino acid residues (Fig. 8). A site after a termination codon, a splice removes the remainder of intron F (designated Fb in Fig. 8) utilizing the same 3' splice acceptor site at the 5' end of exon VII as in the normally spliced transcript. Thus, the 3' end of the alternatively spliced transcript is derived from exon VII as in XCl, but this region becomes a noncoding region in XCl, since it is preceded by a translation termination signal.

The effect of the alternative splice is to replace the carboxyl-terminal 40 amino acid residues of the C subunit (Fig. 8, exon VII) with a new domain of 56 residues (Fig. 8, exon N) generating a protein of 374 amino acids. The C subunit containing the novel open-reading frame of XCl is designated CeCAT'. The C subunit containing the open-reading frame of XCl is highly homologous to mammalian C subunits and is designated CeCATa. The novel exon within intron F is denoted exon N and is bordered by introns F'a and F'b (Figs. 5, 8, and 10A).
Since most of the sequence of the cDNA insert subcloned in labeled insert from a subclone of XC3, pCAT6 (Fig. 2A). pCAT6 (Fig. 2A) is present in both the CeCATa and the alternate 3' splice acceptor site within intron F is utilized, intron F'b is preserved within the cDNA. Since this region again generating exon N. However, the splice between the 3' end of exon N and exon VII is not present in XC2, so that cDNA encodes a full-length CeCATa' subunit. Therefore, tively spliced exon N), the lack of this splice is silent, and the transcript (due to the termination signal within the alterna-

Occurrence of Multiple Sizes of Catalytic Subunit mRNAs—The cDNA clone XC2 (Fig. 2A) is similar to XC3, in that Northern Analysis of C. elegans RNA Demonstrates the... one major 2.5 kb mRNA as well as two larger, but much less abundant mRNAs are detected by hybridization with the 32P-labeled insert from pCAT6. One or more of these three bands could be due to alternative polyadenylation sites or transcription initiation sites, but they might also represent the form of mRNA with a 261-bp insertion corresponding to ACC3 and the form of the mRNA cloned as cDNA in XC2.

PCR Amplification Confirms the Existence of CeCATa' Transcripts in C. elegans—To confirm that CeCATa' was not an unusual cloning artifact, its existence was verified by an independent method. Since the abundance of the larger RNAs detected by Northern analysis appeared to be very low (Fig. 9), we used the more sensitive PCR amplification method. The design of the experiment is shown in Fig. 10A: total RNA was reverse-transcribed using an anti-sense oligonucleotide primer for a region in exon VII common to both the CeCATa and CeCATa' mRNAs (Fig. 10A, primer 64). Reverse-transcribed, single-stranded cDNA products were then amplified using a common upstream oligonucleotide primer from exon V (Fig. 10A, primer 26) and 3' oligonucleotides (primer 14 or primer 65, Fig. 10A) upstream from the oligonucleotide primer (Fig. 10A, primer 64) used for reverse transcription. One predominant product with a size of 677 bp was expected from the PCR reaction using the common 3' anti-sense oligonucleotide primer 14 (Fig. 10A). Lane 5 of Fig. 10B shows that a product of 677 bp was detected by agarose gel electrophoresis. A product of 778 bp was anticipated from the PCR reaction using the 3' anti-sense oligonucleotide specific for the alternatively-spliced exon N (Fig. 10A, primer 65). Lane 6 shows that a cDNA of the expected size was amplified. To establish the nature of these products, agarose gels were dried and hybridized with internal oligonucleotides from a common region (Fig. 10A, oligonucleotide 66) or from the exon specific for CeCATcJ (Fig. 10A, oligonucleotide 66). The autoradiogram in Fig. 10B shows that the common 3' probe hybrizes with the PCR products from both sets of primers (Fig. 10B, lanes 1 and 2). However, the probe specific for the alternative exon (oligonucleotide 66, Fig. 10A) detects only the 778-bp PCR product (Fig. 10B, lane 4). Thus, this cDNA is an amplification product of CeCATa' transcripts present in C. elegans RNA. The product obtained from oligonucleotide primer 14 does not hybridize with the alternative exon-specific probe (Fig. 10B, lane 3) and is therefore derived from CeCATa, as suggested by its size and its hybridization with oligonucleotide 65.

The Predicted Carboxyl Terminus of CeCATa' Generates a Novel Form of C Subunit—The translation of ACC3 cDNA predicts that CeCATa' is a protein of 374 amino acids, 16 residues larger than CeCATa. Protein comparison programs revealed little similarity between the two forms of carboxyl termini. The region of CeCATa that is replaced in CeCATa', as suggested by its size and its hybridization with oligonucleotide 65.

FIG. 8. cDNA and derived amino acid sequence for the carboxyl termini of two forms of the C. elegans C subunit. DNA sequence analysis was performed as described under "Experimental Procedures." Both forms of C. elegans C are identical through amino acid residue 318. In CeCATa', intron F (between exon N + intron F'b) is removed by splicing and the carboxyl-terminal 41 residues are encoded by exon VII; in CeCATa' introns F'a and F'b are spliced out, thereby producing exon N. Exon N encodes a novel 57-residue carboxyl-terminal sequence and introduces a new tran-...
C. elegans Catalytic Subunit

Amino acid residues 325–339 of CeCATα, whereas CeCAT-A' (VDGPADTRHFVVEEVQ) includes amino acids 341–355 in CeCATα'. The sizes and composition of the peptides were based on the recommendations of Lerner et al. (46, 47) for CeCATJ. The sizes and composition of the peptides were based on the recommendations of Lerner et al. (46, 47) for CeCATJ. The peptides were coupled to the carrier protein thyroglobulin with glutaraldehyde and were injected into rabbits to produce high affinity sera.

Anti-CeCAT-A IgGs bind to three polypeptides with apparent Mr values of 39,000–41,000 on Western blots (Fig. 11, lanes 1 and 2) of C. elegans cytosolic and particulate proteins. The synthetic peptide CeCAT-A (30 µg/ml) inhibits the formation of the three antigen-antibody complexes (Fig. 11, lanes 3 and 4). The same antibody binds a single 40-kDa C subunit in murine erythroleukemic cell cytosol (Fig. 11, lane 7). Anti-CeCAT-A serum precipitates >95% of the phosphotransferase in C. elegans cytosol (data not shown). Anti-CeCAT-A' serum binds with one 41-kDa protein to yield a band of faint intensity (Fig. 11, lane 5), as anticipated for a nonabundant C isoform. The binding is inhibited by addition of peptide CeCAT-A' (Fig. 11, lane 6). Thus, the alternative α' isoform of C. elegans C is detected as a nonabundant component of the cytosol.

Developmental Regulation of C Subunit Expression—C. elegans embryos were isolated and synchronized by a standard procedure (see "Experimental Procedures"). Nematodes at various stages of development were harvested and disrupted by sonication. The relative levels of the 39–41-kDa C subunits were monitored by Western blot analysis (Fig. 12A) using anti-CeCAT-A serum. Early and late (prehatching) embryos exhibit low levels of C subunits (Fig. 12A, lanes E and PH (4 h prior to hatching)). As the embryos hatch (Fig. 12A, lane H), a burst of C subunit expression is observed. Integrating densitometry of the autoradiograms and direct assays of C subunits were detected in both cytosolic and particulate fractions respectively, during the period of transition from embryo to the first larval stage (Fig. 12A, lane H). Elevated levels of the two larger forms of C are found in the L1 to L4 larval stages. Typical data from L3 animals are shown in Fig. 12A. C subunits were detected in both cytosolic and particulate fractions, thereby paralleling the distribution of R subunits. As C. elegans matures into the adult stages, cytosolic C apparently vanishes, whereas particulate C is diminished (Fig. 12A, lanes A and EL). In order to determine whether the decline in C levels was due to endogenous physiological degradation of C subunits or a post-sonication artifact, adult (A), egg-laying (EL), and hatched (H) C. elegans were pelleted, immediately dissolved in buffer containing 2% SDS, 1% mercaptoethanol by heating for 5 min at 100 °C, and the samples were then subjected to electrophoresis and Western analysis.

Under these conditions the content of C was essentially the same in the three samples (Fig. 12B), suggesting that proteases expressed in adult nematodes degrade C subunits during fractionation procedures despite the presence of various protease inhibitors in the homogenization buffer. The levels of C in embryos and pre-hatched worms were not affected by this treatment.

Chromosomal Location of the C. elegans Catalytic Subunit Gene—The catalytic subunit gene clones XCG1 and XCG2 (Fig. 2B) were sent to Drs. A. Coulson, J. Sulston, and D. Albertson (Medical Research Council Laboratory of Molecular Biology, Cambridge, Great Britain) for assignment of C to a location on the physical map of the C. elegans genome. XCG1 and XCG2 were "fingerprinted" by the method of Coulson et al. (6) and were found to overlap cosmids within a contig on the extreme end of the right arm of chromosome 1. The catalytic subunit gene, named kin-1, is flanked by unc-54 and let-50.

DISCUSSION

Cytosolic and particulate fractions isolated from asynchronous populations of C. elegans contain levels of cAMP-dependent protein kinase activity that are similar to the levels of cAMP-activated phosphotransferase activity expressed in mammalian tissues (Table I). These results are consistent with a possible role for the enzyme in signal transduction in C. elegans. The preponderance of particulate cAMP-dependant protein kinase in the nematode (65% of total units) parallels the distribution of cAMP-dependent protein kinase
in mammalian brain (40) but differs from the more typical cytosolic localization (≥90% cytosolic) of the enzyme in other mammalian cells and tissues (40, 48).

Expression of *C. elegans* C is developmentally regulated. The specific activity and amount of C subunit protein is low during embryogenesis (Fig. 12A). However, C content increases -15-fold when the nematodes hatch (Fig. 12A). This level declines slightly as the animals progress through further stages of development. The data do not preclude a role for C in embryogenesis, but they suggest that an increased concentration of cAMP-dependent protein kinase might be needed for the post-embryonic progression through the larval L1, L2, L3, and L4 stages, as well as the development of reproductive adult nematodes. It is apparent that the precursors rise in C content and activity correlates with the time at which the animals must acquire the capability of responding to and transducing environmental (e.g. chemosensory, mechanosensory, food, etc.) signals. It is also possible that the increase in C is essential for the hatching process itself. The relative abundance of C mRNA varies only slightly during nematode development. Typical measurements are presented in Table III. Thus, developmental regulation is exerted at the translational and/or post-translational levels.

Two types of cDNA coding *C. elegans* C subunits have been cloned and sequenced. Isoform CeCATα cDNA hybridizes with a major 2.5-kb mRNA and minor mRNAs that are 2.7 and 2.9 kb in length (Fig. 9) and yields a derived amino acid sequence that is 82% identical with the mouse Ca subunit (Fig. 4). The predicted CeCATα polypeptide contains conserved sequences for components of the ATP binding and catalytic sites in the core catalytic sequence (residues 40–300) defined by Hanks et al. (41) for the prototypic bovine Ca subunit. The principal regions of divergence between the CeCATα and bovine Ca sequences occur near the amino and carboxyl termini (Fig. 4). Overall, C subunit sequences are highly conserved in two distantly related branches of evolution. In contrast, a *C. elegans* R subunit that we recently cloned and sequenced is only 58% identical with a mammalian R subunit. Moreover, *C. elegans* R contains a 100-residue segment that is highly divergent from the corresponding region in all mammalian R subunit isoforms.

Southern blot analyses indicated that *C. elegans* C is the product of a unique gene. Even under low stringency conditions, using pCAT-1 and pCAT-2 cDNA inserts (Fig. 2A) as probes, no restriction fragments other than those that represent the C gene in Fig. 1 were detected (data not shown). The 20-kb *C. elegans* C gene is similar in size to the murine Ca gene (50). However, its intron/exon organization (Fig. 5) differs. The murine Ca gene contains 10 exons and the conserved sequences GXXGXXG-17-K, DLKPHN, DFG-19-APE, and DWALG are in exons III, VI, VII, and VIII, respectively; *C. elegans* C coding sequence is interrupted by only six introns and contains GXXGXXG-17-K in exon III, DLKPHN in exon V, and both DFG-19-APE and DWALG in exon VI.

The C subunit gene is uncharacteristically large for *C. elegans*. This is due to the presence of two introns of 8 and 7.7 kb in the 5′ end of the gene (Table II, Fig. 5), whereas 48–52 bp is the average intron size reported for a number of *C. elegans* genes (43). Whether there is any significance to the large introns in an organism with a small genome size is unclear, as is the asymmetric distribution of large introns. It is possible to speculate that genes that are highly expressed in *C. elegans* (43) have small introns. Alternatively, large introns in the 5′ end of the gene might play a role in C gene regulation, either by affecting splicing efficiency or by influencing the rate of transcription or initiation.

The significance of the final 1034-bp intron is more apparent. In the CeCATα isoform (Figs. 8 and 10A) intron F is spliced out of the pre-mRNA, thereby generating a primary sequence that is colinear with the previously characterized bovine and murine Ca cDNAs (16, 50, 51). ACC3 which encodes CeCATα includes the same sequence as ACC1 but has an insertion of 261 bp derived from an internal region of intron F (Figs. 5 and 8). The inserted sequence encodes an open-reading frame of 57 amino acid residues, in register with the upstream (5′) C subunit coding sequence. After residue 57 in the insert there is a termination codon, so that the 3′ end of this alternatively spliced sequence and all of exon VII in CeCATα become noncoding. The effect of the alternative splice is to replace the usual carboxyl-terminus of the C subunit with a distinct 56 residue segment (Fig. 8).

Could this form of C cDNA be a cloning artifact? For example, artifacts occur when cDNA is reverse-transcribed from nuclear precursor species which contain introns. The mechanism of splicing currently supported by substantial biochemical evidence (52) involves the en bloc removal of introns. Therefore, splicing intermediates contain entire introns, not parts of introns. ACC3 could not have been derived in such a fashion, however, since it contains a fragment of intron F (exon N, Fig. 8). Furthermore, this internal fragment of intron F (exon N, Figs. 8 and 10A) is bounded by consensus *C. elegans* splice acceptor and donor sequences.

In addition to isolating cDNA clones containing exon N, PCR studies support the presence of CeCATα′ transcripts in the *C. elegans* poly(A)+ RNA pool. We reverse-transcribed mRNA with an oligonucleotide primer (Fig. 10A) that was designed from a region of intron F that could only be present in the mRNA if (1) exon N were spliced into a transcript as in CeCATα or (2) there was a splicing intermediate in which intron F is not yet removed. After amplification of the cDNA with Taq DNA polymerase and an upstream primer from exon V (see Fig. 10A), the putative CeCATα′ transcript should yield a 778-bp cDNA product, whereas a nuclear precursor containing unspliced intron F would yield a 1550-bp cDNA. A 778-bp product was detected (Fig. 10B), and this DNA hybridized with an independent end labeled oligonucleotide that was also designed from exon N (Fig. 10).
producing new carboxyl-terminal sequences and transforming the final exon into 3' untranslated sequence (52). This type for a novel C. elegans Ser/Thr protein kinase.

since we have observed the same phenomenon in cDNA clones genes encoding myosin heavy chain, alkali myosin light chain, in C. elegans. The IgGs detected a nonabundant protein that exhibits an apparent Mr expected for a C subunit (Fig. 11). The binding of the antibodies was inhibited by an excess of peptide A'.

Aplysia, alternative splicing of C subunit pre-mRNA also generates two phosphotransferases (53). This is due to the mutually exclusive utilization of two forms of an exon encoding amino acid residues 142–183. The net effect of alternate splicing in Aplysia is the replacement of the 10 of the 42 amino acid residues encoded by the exon cassette (53). The sequences involved in ATP binding and catalysis in mammalian C subunits are conserved in the two C isoforms of Aplysia. Both isoforms are expressed in Aplysia neural tissue but the isoform that most closely resembles mammalian C accounts for >80% of the phosphotransferase mRNA. It is not yet known whether the two C isoforms in Aplysia exhibit differences in kinetic properties, substrate specificity, or subcellular localization.

The difference between CeCATα and CeCATα' is striking: the final 40 amino acid residues in the C. elegans isoform (CeCATα) that closely resembles mammalian C subunits are replaced by a unique segment of 56 carboxyl-terminal residues in CeCATα' (Fig. 8). Both isoforms are identical in their core catalytic domains (residues 40–310). Since the carboxyl-terminal region is variable in related Ser/Thr protein kinases and little is known about structure/function relationships in the carboxyl terminus of C subunits it is not possible to speculate on the effects of the novel 56-residue segment on the enzymic properties or localization of CeCATα'.

The occurrence of alternative splicing in a carboxyl-terminal exon that generates isoforms and transforms a portion of coding sequence into nontranslated 3' sequence is a rare event, but is not without precedent. First, the presence of single 5' donor sites that may be spliced to two or more 3' acceptor sites have been reported for fibronectin, adenovirus, and SV40 pre-mRNAs (52). As is the case in C. elegans, the utilization of such alternative sites blurs the usual distinction between "exon" and "intron" sequences. Second, Drosophila genes encoding myosin heavy chain, alkali myosin light chain, tropomyosin I, and P element contain alternative exon sequences that introduce new translation stop signals, thereby producing new carboxyl terminal sequences and transforming the final exon into 3'-untranslated sequence (52). This type of splicing pathway might be used frequently in C. elegans since we have observed the same phenomenon in cDNA clones for a novel C. elegans Ser/Thr protein kinase.

The 5' end of the C. elegans C mRNA appears to be derived from trans-splicing of the 22-bp of a 94-bp "leader" RNA to a 3' acceptor site in the C subunit primary transcript (44). The leader sequences are spliced to approximately 10% of C. elegans mRNAs, but no function has yet been ascribed to them (45). The C. elegans regulatory subunit mRNA is not trans-spliced, thereby providing an interesting example of a multisubunit enzyme whose cognate pre-mRNAs are differentially processed. The trans-splicing mechanism appears to be similar to that utilized for cis-splicing (53). Experimentally, the trans-splicing event makes it difficult to define the transcription start site of a gene. Because the 5' end of the cDNA is not derived from the gene, it would be necessary to isolate cDNAs derived from nuclear precursor RNAs that have not yet been trans-spliced, in order to determine the 5' end of the transcription unit. Thus, it is especially difficult to define the start site of transcription for genes that (1) are expressed at low levels, such that nuclear precursors are difficult to clone, and (2) have multiple start sites and no TATA box, so that the regulatory region of the gene is difficult to recognize. The mouse Ca and CD genes have no TATA or CAAT boxes and exhibit multiple initiation sites (50). We have sequenced >500 bp upstream of the trans-splicing site in the C gene and find no characteristic regulatory regions. Therefore we cannot define the transcription start site or upstream control regions for the C. elegans C gene at present.

Additional diversity in C. elegans C subunits was discovered when Western blots were probed with anti-CeCAT-A peptide IgGs (Figs. 11 and 12). Typically, the antibodies bind three polypeptides with apparent Mr values of 39,000–41,000 in C. elegans cytosol, but complex only one 40-kDa protein in cytosol) from a mammalian cell line that expressed Ca (Fig. 11). In view of the Southern blot data indicating a single C gene (Fig. 1) and the specificity of the antibody (see "Results" and Fig. 11), it appears that multiple forms of C might be a consequence of post-translational modifications (e.g. phosphorylation). Unlike mammalian C subunits, C. elegans C apparently does not undergo amino terminal myristoylation because the initiator Met is not followed by the requisite Gly residue (Figs. 3 and 4 and Ref. 54). It will be necessary to purify and characterize the multiple forms of C biochemically and by labeling and immunoprecipitation to determine the molecular basis for this additional level of diversity in C subunits. Previous work on the C subunit in other systems has pointed to certain sites that when altered by site-directed mutagenesis lead to a perturbation of C function. For example, changing the Thr at position 197 to an Ala generated a yeast C subunit that is not inhibited by R subunits and is therefore constitutively active (42). Subunits mutated in such a fashion are dominant relative to wild-type subunits. The availability of gene, cosmid, and cDNA clones of C. elegans C subunits, the potential for constructing transgenic animals and the cloning of C. elegans regulatable and cell-specific promoters will allow the manipulation of cAMP-dependent protein kinase expression in the nematode. If cAMP-dependent protein kinase is a component of developmental and/or neurotransmitter signal transduction pathways in C. elegans a change in phenotype might be expected in such trangenic animals. It would then be possible to elucidate the components of these pathways downstream from cAMP dependent protein kinase by isolating animals that suppress or enhance that phenotype.

Acknowledgments—We thank Dr. Philip Anderson and Claudia Cummins (Department of Genetics, University of Wisconsin) for generously providing a cDNA probe that hybridizes with the myosin light chain mRNAs designated mlc-1 and mlc-2 (Ref. 49). We also wish to acknowledge Bridget Hickey for expert secretarial services.

REFERENCES
**E. coli** Catalytic Subunit

Natl. Acad. Sci. U. S. A. 89, 2172-2176


34. Hall, P. B., Marder, E., and Lis, J. (1985) EMBO J. 4, 2727-2734


C. elegans Catalytic Subunit

C. elegans Catalytic Subunit 6905

DNA Sequencing

Restriction digestion of DNA from C. elegans genomic DNA with HindIII provided a 6.9 kb fragment corresponding to the C. elegans gene. Yeast transformation revealed that the HindIII fragment appeared to be the catalytic subunit of C. elegans. The DNA sequence was then determined by the dideoxynucleotide chain termination procedure of Sanger (1977) using a Pharmacia automated DNA sequencer.

DNA Sequence Analysis

Restriction fragments of DNA from C. elegans genomic DNA were subcloned into E. coli and sequenced using the dideoxynucleotide chain termination procedure of Sanger (1977) using a Pharmacia automated DNA sequencer.

Yeast Transformation

The DNA sequence was then determined by the dideoxynucleotide chain termination procedure of Sanger (1977) using a Pharmacia automated DNA sequencer.

Antisense Oligonucleotide

An antisense oligonucleotide was synthesized and used as a probe for in situ hybridization. The antisense oligonucleotide was complementary to the 3' end of the C. elegans catalytic subunit cDNA.

Northern Blot Analysis

A Northern blot analysis was performed using total RNA from C. elegans embryos and adult animals. The results showed that the catalytic subunit mRNA is expressed in both embryonic and adult tissues.

Table 1

<table>
<thead>
<tr>
<th>Source of Enzyme</th>
<th>Specific Activity (U/mg of protein)</th>
<th>Field Stimulation by 5 mM CAMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. elegans exoyld</td>
<td>2.4</td>
<td>2.8</td>
</tr>
<tr>
<td>C. elegans inoxyl</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>specific activity</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>bovine brain exoyld</td>
<td>6.5</td>
<td>6.5</td>
</tr>
<tr>
<td>rabbit brain exoyld</td>
<td>6.5</td>
<td>6.5</td>
</tr>
<tr>
<td>mouse muscle exoyld</td>
<td>4.4</td>
<td>4.4</td>
</tr>
<tr>
<td>bovine brain membranes</td>
<td>8.5</td>
<td>8.5</td>
</tr>
</tbody>
</table>

* Averages were performed as described by Rabinovitch et al. 1981 using histone H2B in a substrate. Specific activity is reported as the activity measured in the presence of 5 mM CAMP minus the activity observed in the absence of cyclic AMP.
Figure A. Sizes and partial restriction maps of cDNA and genomic DNA clones for the C. elegans C subunit. (A) The diagram shows the sizes, alignments, and partial restriction maps of cDNA clones (35C, 48A) obtained by screening two cDNA libraries as described under Experimental Procedures and in the text. Subclones (pCA7-pCA27) derived from the primary cDNA clones are shown below the primary clones. (B) The diagram presents the sizes, alignments, and partial restriction maps of genomic DNA fragments that were characterized in order to determine the intron/exon organization of the C. elegans C gene. The entire C gene was generated from a subclone (not depicted) that was sequenced. The positions of sites cloned by EcoRI (B), Hind III (H) and Hind I (K) are shown.

Figure B. Determination of the origin of the sequence at the 5' end of C. elegans C subunit. The 5' ends of the sequences shown are on the left.

Figure C. Mapping the 5' end of C. elegans C subunit mRNA by S1 nuclease digestion and primer extension synthesis. C. elegans poly A-RNA (20 μg) was hybridized with 0.4 ng of the 5'-labeled, 40-mer oligonucleotide complementary to CR25 (see Fig. 4 and Experimental Procedures). The resulting complexes were digested with 10 μg of S1 nuclease (23) and the products were separated on 5% polyacrylamide (60:20 M) gels as described under Experimental Procedures. A EcoRI restriction enzyme digestion was performed on the 50-mer oligonucleotide to generate 39 and 38-mer fragments (lane 3). Lane 1 depicts the probe (60 bp) and the principal protected products (56-58 bp) obtained. Then 10 μg of total RNA was substituted for poly A-RNA (lane 2). The protected products were not observed (data not shown). For primer extension synthesis 2 μg of the 5'-labeled 30-mer oligonucleotide complementary to CR25 (Fig. 4 and Experimental Procedures) was hybridized with 10 μg of poly A-RNA as outlined under Experimental Procedures. After extension of the labeled primer with reverse transcriptase the 15'-labeled products were characterized by electrophoresis in the polyacrylamide gel used for S1 nuclease analysis (see Experimental Procedures). The autoradiogram shows the extended products in lane 4. Undetected products were observed in lane 5. The gel was calibrated with four λ Hind III fragments (25 pg each).
Figure 3. Northern gel analysis of C. elegans C subunit mRNA. C. elegans poly A+ mRNA was prepared, electrophoresed, transferred to a nitrocellulose sheet, and hybridized with the 32P-labeled cDNA insert from pCET (Fig. 2A) as described under Experimental Procedures. The gel was calibrated by running a standard RNA ladder in a parallel lane and staining with ethidium bromide. The size of the major C subunit mRNA is given; the positions of the minor mRNAs are indicated with dots.

Figure 4A. Amplification and identification of Cα1-α4 transcripts. A schematic diagram showing the cDNA coding sequence corresponding to Cα1-α4 and CαT-α5 mRNAs is presented in 8B. The positions of the shared inverse transcription (CI1) and specific amplification (CI4) primers and the second specific primer (CI6) are indicated. Some of the expected amplification products are shown for the two combinations of primers indicated on the right. The 32P-labeled oligonucleotide was hybridized with each of the amplified cDNA; whereas 32P-labeled CI1 is an exon 6-specific probe that will hybridize only to sequences 3′ to either of the CI1 primers, the CI6 probe hybridizes only to sequences 3′ to both the CI6 primers. The sequences of the primers and conditions for renaturation and hybridization of the oligonucleotides are given under Experimental Procedures. A 0.5-μl aliquot (300 ng of DNA) was reverse-transcribed and then amplified by the PCR reaction under the primer described above and the conditions indicated under Experimental Procedures. Details of PCR reaction conditions are given in 0.5 μl of Cα1-α4 DNA product. 0.5 μl of Cα1-α4 DNA product was applied to lanes 1, 3, and 5 of an agarose gel. PCR products obtained with CI1 (CI), CI6 (CI), CI4 (CI4), CI1 (CI1), CI6 (CI6), and CI4 (CI4) oligonucleotides were applied to lanes 2, 4, and 6. After electrophoresis, lanes 2, 4, and 6 were stained with ethidium bromide. Lanes 1 and 3 were stained with ethidium bromide. Lanes 2 and 6 were dried and probed with 32P-labeled CI6. Autoradiograms are shown for lanes 1-6.

R E Gross, S Bagchi, X Lu and C S Rubin


Access the most updated version of this article at [http://www.jbc.org/content/265/12/6896](http://www.jbc.org/content/265/12/6896)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/265/12/6896.full.html#ref-list-1](http://www.jbc.org/content/265/12/6896.full.html#ref-list-1)