

# The *C. elegans* Cell Death Gene *ced-3* Encodes a Protein Similar to Mammalian Interleukin-1 $\beta$ -Converting Enzyme

Junying Yuan,\*<sup>††</sup> Shai Shaham,<sup>†</sup>  
Stephane Ledoux,<sup>‡§</sup> Hilary M. Ellis,<sup>†||</sup>  
and H. Robert Horvitz<sup>†</sup>

\*Program of Neurosciences

Harvard Medical School

Boston, Massachusetts 02115

<sup>†</sup>Howard Hughes Medical Institute

Department of Biology

Massachusetts Institute of Technology

Cambridge, Massachusetts 02139

<sup>‡</sup>Cardiovascular Research Center

Massachusetts General Hospital

Charlestown, Massachusetts 02129

## Summary

We have cloned the *C. elegans* cell death gene *ced-3*. A *ced-3* transcript is most abundant during embryogenesis, the stage during which most programmed cell deaths occur. The predicted CED-3 protein shows similarity to human and murine interleukin-1 $\beta$ -converting enzyme and to the product of the mouse *nedd-2* gene, which is expressed in the embryonic brain. The sequences of 12 *ced-3* mutations as well as the sequences of *ced-3* genes from two related nematode species identify sites of potential functional importance. We propose that the CED-3 protein acts as a cysteine protease in the initiation of programmed cell death in *C. elegans* and that cysteine proteases also function in programmed cell death in mammals.

## Introduction

Cell death occurs as a normal aspect of animal development as well as in tissue homeostasis and aging (Glucksmann, 1950; Ellis et al., 1991a). Naturally occurring or programmed cell death can act to regulate cell number, to facilitate morphogenesis, to remove harmful or otherwise abnormal cells, and to eliminate cells that have already performed their functions. In many cases, gene expression within dying cells is thought to be required for these cells to die, since the cell death process can be blocked by inhibitors of RNA and protein synthesis (Stanisic et al., 1978; Cohen and Duke, 1984; Martin et al., 1988).

During the development of the nematode *Caenorhabditis elegans*, 131 cells undergo programmed cell death (Sulston and Horvitz, 1977; Sulston et al., 1983). Fourteen genes have been identified that function in different steps of the genetic pathway of programmed cell death in *C. elegans* (Hedgecock et al., 1983; Ellis and Horvitz, 1986,

1991; Ellis et al., 1991b; Hengartner et al., 1992; reviewed by Ellis et al., 1991a). Two of these genes, *ced-3* and *ced-4*, play essential roles in either the initiation or execution of the cell death program, since recessive mutations in these genes prevent almost all of the cell deaths that normally occur during *C. elegans* development. Genetic mosaic analysis indicates that *ced-3* and *ced-4* most likely function within cells that die or within their close relatives to cause cell death (Yuan and Horvitz, 1990). The *ced-4* gene encodes a novel protein that is expressed primarily during embryogenesis, the period during which most programmed cell deaths occur (Yuan and Horvitz, 1992).

To understand how the *ced-3* gene acts to cause cell death, we have cloned this gene. As deduced from the sequence of a *ced-3* cDNA clone, the CED-3 protein is 503 amino acids in length and contains a serine-rich middle region of about 100 amino acids. We compared the sequences of the CED-3 protein of *C. elegans* with the inferred CED-3 protein sequences from the related nematode species *C. briggsae* and *C. vulgaris*. This comparison revealed that the carboxy-terminal portions of these proteins are most conserved. The non-serine-rich portions of the CED-3 protein are similar to human interleukin-1 $\beta$  (IL-1 $\beta$ )-converting enzyme (ICE), a cysteine protease that can cleave the inactive 31 kd precursor of IL-1 $\beta$  to generate the active cytokine (Cerretti et al., 1992; Thornberry et al., 1992). In addition, the carboxy-terminal portions of both the CED-3 and ICE proteins are similar to the mouse *Nedd-2* protein, which is encoded by a messenger RNA (mRNA) expressed during mouse embryonic brain development and is down-regulated in adult brain (Kumar et al., 1992). We suggest that CED-3 acts as a cysteine protease in controlling the onset of programmed cell death in *C. elegans* and that members of the *ced-3*/ICE/*nedd-2* gene family might function in programmed cell death in vertebrates.

## Results

### *ced-3* Is Not Essential for Viability

All previously described *ced-3* alleles were isolated in screens designed to detect viable mutants in which programmed cell death did not occur (Ellis and Horvitz, 1986). Such screens might systematically have missed classes of *ced-3* mutations that result in inviability. Since animals of the genotype *ced-3*/deficiency are viable (Ellis and Horvitz, 1986), we designed a screen that would allow us to isolate recessive lethal alleles of *ced-3*, if such alleles could exist (see Experimental Procedures). We obtained four new *ced-3* alleles (*n1163*, *n1164*, *n1165*, and *n1286*) in this way. All four of these mutants are viable as homozygotes. These alleles were isolated at a frequency of about 1 in 2500 mutagenized haploid genomes, approximately the frequency expected for the generation of loss-of-function mutations in an average *C. elegans* gene (Brenner, 1974; Meneely and Herman, 1979; Greenwald and Horvitz, 1980).

<sup>§</sup>Present address: Neuroimmunology Laboratory, Montreal Neurological Institute, Montreal, Quebec H3A 2B4, Canada.

<sup>||</sup>Present address: Department of Biology, Emory University, Atlanta, Georgia 30322.

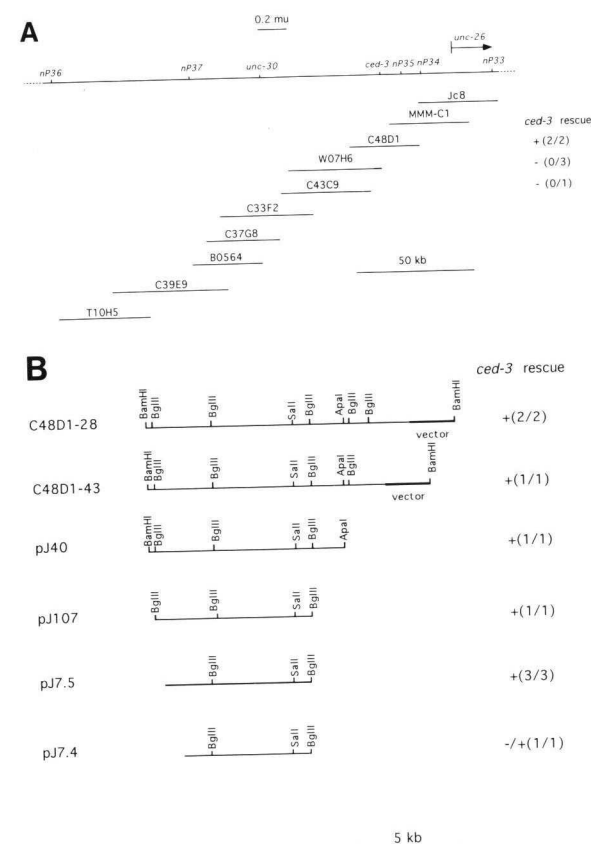


Figure 1. Genetic and Physical Maps of the *ced-3* Region of Chromosome IV

(A) Alignment of the genetic and physical maps in the *ced-3* region. *nP33*, *nP34*, *nP35*, *nP36*, and *nP37* are RFLPs between the Bristol N2 and Bergerac EM1002 wild-type *C. elegans* strains. The three cosmid clones C43C9, W07H6, and C48D1 were tested for their abilities to rescue the *Ced* phenotype of *ced-3(n717)* animals. The ability of each cosmid clone to rescue *ced-3* mutants and the fraction of independently obtained transgenic lines that were rescued are indicated on the right. Symbols: plus, rescue; minus, no rescue. See text for data. The results indicate that *ced-3* is contained in cosmid C48D1.

(B) A restriction map of subclones of cosmid C48D1. Bold lines represent cosmid vector sequences. Subclones were assayed for rescue of the *Ced* phenotype of *ced-3(n717)* animals as in (A). Symbols: plus, rescue; minus, no rescue; minus/plus, weak rescue (fewer than five corpses on average). The numbers in parentheses indicate the fraction of independently obtained transgenic lines that were rescued. The smallest fragment that fully rescued the *Ced-3* mutant phenotype was the 7.5 kb pJ7.5 subclone.

These observations suggest that animals that lack *ced-3* gene activity are viable. Supporting this hypothesis, we have shown by molecular analysis that three *ced-3* mutations are nonsense mutations that seem likely to eliminate *ced-3* activity (see below). Based upon these considerations, we conclude that *ced-3* gene activity is not essential for viability.

#### *ced-3* Is Contained within a 7.5 kb Genomic Fragment

To clone the *ced-3* gene, we used the approach of Ruvkun et al. (1988). In brief (for further details see Experimental Procedures), we identified a 5.1 kb EcoRI restriction frag-

ment that contained Tc1 (Emmons et al., 1983), that was present in the *C. elegans* Bristol strain N2 but not in the *C. elegans* Bergerac strain EM1002, and that was closely linked to *ced-3*. We named this restriction fragment length polymorphism (RFLP) *nP35*. Using Tc1 to probe Southern blots of cosmids derived from N2 genomic DNA and known to contain Tc1 (G. Ruvkun, personal communication), we identified two cosmids (MMM-C1 and MMM-C9) that contained this 5.1 kb EcoRI fragment. These cosmids overlapped an existing cosmid contig that had been defined as part of the *C. elegans* genome project (Coulson et al., 1986, 1988; A. J. Coulson et al., personal communication). We used cosmids from this contig to identify four additional Bristol-Bergerac RFLPs (*nP33*, *nP34*, *nP36*, and *nP37*). By mapping these RFLPs between the Bristol and Bergerac strains with respect to the genes *unc-30*, *ced-3*, and *unc-26*, we oriented the contig with respect to the genetic map. These experiments narrowed the region containing the *ced-3* gene to an interval spanned by the three cosmids C48D1, W07H6, and C43C9 (Figure 1A).

These three cosmids were microinjected (Fire, 1986; J. Sulston, personal communication) into *ced-3* mutant animals to test for rescue of the mutant phenotype. Specifically, a candidate cosmid and cosmid C14G10, which contains the wild-type *unc-31* gene (R. Hoskins, personal communication), were coinjected into *ced-1(e1735); unc-31(e928) ced-3(n717)* hermaphrodites, and non-Unc progeny were isolated and observed to see whether the non-Unc phenotype was transmitted to the next generation, thus establishing a line of transgenic animals. Young first larval stage (L1) progeny of such transgenic lines were examined for the presence of cell deaths using Nomarski optics to see whether the *Ced-3* phenotype was rescued. Cosmid C14G10 alone does not confer wild-type *ced-3* activity when injected into a *ced-3* mutant (data not shown). *ced-1* was used to facilitate scoring of the *Ced-3* phenotype (see Experimental Procedures), and *unc-31* was used as a marker for cotransformation (R. Hoskins, personal communication; Kim and Horvitz, 1990).

As indicated in Figure 1A, of the three cosmids tested, only C48D1 rescued the *Ced-3* mutant phenotype. Two non-Unc transgenic lines were obtained, the insertion line *nIs1* and the extrachromosomal line *nEx2*. Both were rescued. Specifically, L1 *ced-1* animals contained an average of 23 cell corpses in the head, and L1 *ced-1; ced-3* animals contained an average of 0.3 cell corpses in the head (Ellis and Horvitz, 1986). By contrast, L1 *ced-1; unc-31 ced-3; nIs1* and L1 *ced-1; unc-31 ced-3; nEx2* animals contained an average of 16.4 ( $n = 20$ ) and 14.5 ( $n = 20$ ) cell corpses in the head, respectively. From these results, we concluded that C48D1 contains the *ced-3* gene.

To locate *ced-3* more precisely within cosmid C48D1, we subcloned this cosmid and tested the subclones for their abilities to rescue the *Ced-3* mutant phenotype (Figure 1B). From these experiments, we localized *ced-3* to a DNA fragment of 7.5 kb (pJ7.5).

#### *ced-3* Is Transcribed Primarily during Embryogenesis and Independently of *ced-4* Function

We used the 7.6 kb pJ107 subclone of C48D1 (Figure 1B)

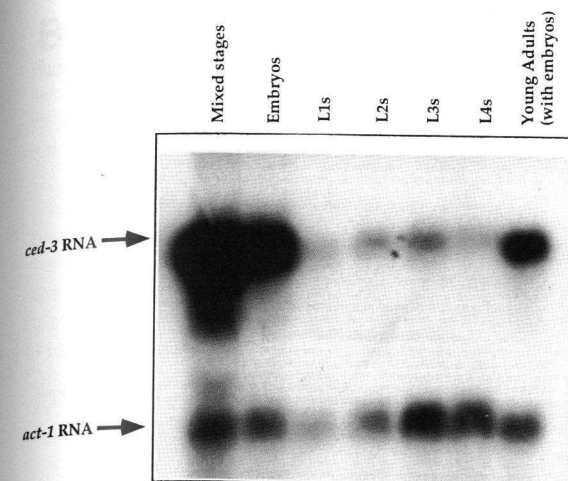


Figure 2. *ced-3* RNA Is Expressed in Embryos

A Northern blot of poly(A)<sup>+</sup> RNA from mixed stages, embryos, L1-L4 larvae, and young adults, probed with the *ced-3* cDNA subclone pJ118 (see Experimental Procedures). pJ118 did not detectably hybridize to RNA derived from *glp-1(q231)* adults (which lack a germline; Austin and Kimble, 1987) (data not shown), suggesting that the *ced-3* RNA detected in the young adults in this experiment was derived from embryos within these animals. The level of RNA in each lane can be estimated based upon hybridization to a control actin 1 probe (Krause and Hirsh, 1984).

to probe a Northern blot of poly(A)<sup>+</sup> RNA derived from the wild-type *C. elegans* strain N2. This probe hybridized to a 2.8 kb transcript (data not shown; also see Figure 2). Although this transcript was present in 11 different ethyl methanesulfonate (EMS)-induced *ced-3* mutant strains (data not shown; the mutant *n1164* was not tested), subsequent analysis revealed that all 11 mutants contain mutations in the genomic DNA that encodes this mRNA (see below), thus establishing this RNA as a *ced-3* transcript. The *ced-3* transcript was most abundant during embryogenesis, when most programmed cell deaths occur (Sulston and Horvitz, 1977; Sulston et al., 1983), and was also detected at later stages (Figure 2).

Since *ced-3* and *ced-4* are both required for programmed cell death in *C. elegans* and since both are highly expressed during embryonic development (Yuan and Horvitz, 1992; see above), it is possible that one of these genes regulates the mRNA levels of the other. Previous studies showed that *ced-3* does not regulate *ced-4* mRNA levels (Yuan and Horvitz, 1992). To determine whether *ced-4* regulates *ced-3* mRNA levels, we probed a Northern blot of RNA prepared from *ced-4* mutant embryos with the *ced-3* cDNA subclone pJ118. The abundance and size of *ced-3* transcript were normal in the *ced-4* mutants *n1162* (glutamine to ochre nonsense mutation at codon 40), *n1416* (Tc4 insertion into exon 5), *n1894* (tryptophan to opal nonsense mutation at codon 401), and *n1920* (G to A mutation in the intron 3 splice donor site) (Figure 3A). None of the *ced-4* mutants tested has detectable *ced-4* RNA (Figure 3B; Yuan and Horvitz, 1992). Thus, *ced-4* does not seem to affect the steady-state levels of *ced-3* mRNA.

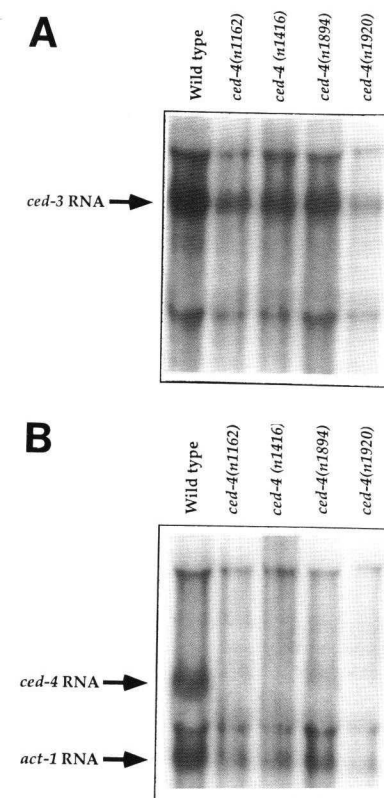


Figure 3. *ced-3* RNA Is Expressed in *ced-4* Mutants

Northern blots of total RNA from wild-type embryos and *ced-4(n1162)*, *ced-4(n1416)*, *ced-4(n1894)*, and *ced-4(n1920)* embryos. Probed with the *ced-3* cDNA subclone pJ118 (see Experimental Procedures) (A) or the *ced-4* probe SK2-1 and a control actin 1 probe (Krause and Hirsh, 1984) (B). Unlabeled bands in (A) and (B) correspond to ribosomal RNA bands. (B) is reprinted from Yuan and Horvitz (1992).

#### *ced-3* cDNA and Genomic Sequences

To isolate *ced-3* cDNA clones, we used the *ced-3* genomic DNA clone pJ40 (see Figure 1B), which rescued the *Ced-3* mutant phenotype, when microinjected into *ced-3* mutant animals, as a probe to screen a cDNA library prepared from the *C. elegans* wild-type strain N2 (Kim and Horvitz, 1990). The 2.5 kb cDNA clone pJ87 was isolated in this way. On Northern blots, pJ87 hybridized to a 2.8 kb transcript, and on Southern blots, it hybridized only to bands to which pJ40 also hybridized (data not shown). Thus, pJ87 was derived from an mRNA transcribed entirely from pJ40. To confirm that pJ87 corresponds to a *ced-3* cDNA clone, we made a frameshift mutation in the unique SalI site in the pJ87 cDNA clone and disrupts the putative *ced-3* open reading frame (see Experimental Procedures; see also Figure 4A). Constructs containing this frameshift mutation failed to rescue the *Ced-3* phenotype when microinjected into *ced-3* mutant animals (six transgenic lines were examined; data not shown), suggesting that *ced-3* activity was eliminated by mutating the putative *ced-3* open reading frame.

We determined the DNA sequence of pJ87 (Figure 4). pJ87 contains an insert of 2482 bp that can encode a protein of 503 amino acids. pJ87 contains 953 bp of 3'

[illegible]

4381 T T C A C C G G A T G C T T C T C T C G A A T A G A T C C A A G C G T A A T T C G C A T T C C A A A A G C T T  
T G C A L L G Y G S S R S S F X A S  
190 200

4391 C T T G A C C A A C T C A M A T A T A T T C C A T A A A A G G A T A G A C T T T C T G A T C C A C C A A C C A  
G P T Q Y I F H E E D M H F V D A P T I  
210 220

4441 T A A G C G T G T T T T C G A C A A A A A A C C A T A C A A A A T T C T C G A G C T G T G G A A A T T  
S R V F D E K E T Y R H F S S P R H C  
230 240

4501 G C C T C A T C A A A A T A A T A A A C A C T T T G A C G A G A T C C C A A C A C G A A G T A C C A A A G C C C  
L I X I H E E F P R Q H P T R H E T K A D  
250 260

4561 A C A A G C A C A A T T A C C A A A T T T T T C A A T C A T G A G C A T A C G G T A T T T C A A A G A C A  
K D N L T S L F R C H E Y V I C K D H  
270 280

4621 A T C T C G C G A A G G T A C G C G A A A T A T A T T A C C A A A C C G A A A A T T T C C A T T T T C G C  
L T G R

Repeat 3

4691 C G A A A A A T T T G C G C C C C G G T C T C G A C A C A C A A T T T T G T T A A T G C A A A A T G T A T A T  
T T T G C A A A A A T A A A T T T G A A C T T C C G C A A A A T G A T T A C T A G T T C G A A A T T T T C  
4741 G T T T T T C C G G C A C A A T A T G T T G T T T T T C T A G T T T T C T A T T A G T A C T A A A A A A  
4801 A C C G T T T T A A T T T C G A C A A T T T T C C G C A C A A A A T T G A T T T T G G C C C A A A A A A  
4861 T T C T G A A A T T T C A A A A T T C C A A A A A C A C A A T T A A A A T T T G A A A A T T G C A A A C  
4921 G G T T T T C A M A T A G A T T A T T T T A A A A C T T T A A A A C C G C T G G A A A A C A A A T A  
4981 A A A T C A A A A C A A C T G C A A A T C A A A T C A A A A T T A C A A A T T T A C A C A A T T T A G A T T  
5041 C A A A T T T T A A A A C A T C A A G A G G T T A A A A A A T T A A C A A T T T T A C A A T T T T A G A T T  
5101 T C A A A A T T T T T T A C A A A A T C G A A A A A A A G A A T G A A A A T G A T T T A A A A A T A T C C  
5161

Repeat 3

5221 A C A G C T T C G A A G A T T T A A A T T A C A G T A C T C C T T A A A G C C A C A C C C C A T T T G C A T T G G  
T G A C C A A A A A T T T T T C T G A G A C C A G T A C C G T A A A T T T T T G T C G A A A A A T T G C A C C A T  
5281 T G A C C A A T A A A C T T C C T A T T C A C C A A A A A T T A A A T T G A A A T T T C G A A A A C C A A A A  
5341 A T T C A A A A A A A A T T C G A A T G A T T T T T T T T T T T T G G C C C A A A A A A A A A A A A  
5401 A A A T C A A T T T T C G A A A A T A C C A A A A A A C C G A A A A A A T T T C C A C G A T T T T C C T  
5461 A A T G T A A A C G A T A T T A A T T C A C G G A A A T G C T C T G A C A A A T C G A G A T T T G C C A A A C  
5521 G H L L T I R D F A E H  
290 300

5581 A C A A T C A C A C G A G A T T C T G G A T A T C T G A T T C A T C A C A C G G A G A A G A A T G T A A  
H S H D G A I L V I L S H E H E M V I  
310 320

5641 T T A T T G A G A T A G A T A T A C G A T A T A C A C A C G A A T A T A G A T C T T C T C A C A C G G  
I G V D D I P I S T H K I Y D L L M A A  
330 340

5701 C A A A T G C T C C C G T C T G C G A A T A A G C C G A A A A T C G T T T T T G C A G G C T T C G A G A G  
N A P R L L A M K P X I V F V Q A C R E  
350 360

5761 G T T C G T T T T A T T T A A T T T A A T A T A A T T T T A A T A A A T T C A T T T C A G A A C T C  
R R

5821 G T G A C A T G A T T C C A G T T T G G A T T C T G C A G C A G A A T T C C T C G A T T T C G T G C T  
D M G F P V L D V D G V F A F L R R G  
370 380

5891 G A T G G A A A C T G A G A C A G C A A T T G T C A A T T T T G T G A T G T T C A G C C G C A A A G T G  
W D N R D G P L F N F L G C V R P Q V Q  
390 400

5941 A G G T G C A A T T A A T T T C T G A A T G A A A T T C C T C A A A A A T C A A A A T G A A T T  
6001 A T T C C A A A A T C C G A T C G A A A A A T T G C A T A T A A T A C G A A A T T G T G A T A A A A G A C

Repeat 4

6061 A A A C C A A T C A C A T C T G C A T C C C G C C A C T T C A T C G G A T T G T T G A A A T G G A G C G A  
T G A A T T C T G A T T G T C G A G T T C G A G T T T C A G T T T A A G A G G A A T T A A A A T G C C T T T C G A  
6121 A A A T A A A A A T T G A T T T T C A T T T T T C G A A A A A T T C G A A A T T A T T A T A T C T T  
6181 A G A G C A A A G C C C C T C C T G T A A A C A T T T T A A A T A G A A T A A T A A A T T T T G C A C A A  
T  
T(n1165)  
410 420

6301 G T T G G A A A A A A A A C C G A C A G C T G A C A T T C G A T T G A T A C G A A C A C A C A G C T C A A  
V W R K K P S G Q A D I L I A Y A T T A Q  
430 440

6361 T A T G T T G T G G A A A A C G T C T C T G T G A T C A T G T P T C A A T T C A A C C C T T G T A A A T G  
Y V S W R M S A R S G W F I Q A V C E V  
450 460

6421 T T C T G A C A C A C G A A A A T A G A T T G T T T G A G C T A G A T G A A T C A A T A A A G A G  
F S T H E K D M D V V E L L T R E V M K K  
470 480

6481 G T C C T G T G G A T T T C A G A C A C A G G A A T C G A A T A T T T G A A A C A G A T C C A G A G A  
V A C G P Q T S Q G S M I L K Q M P E  
490 500

Repeat 5

6541 C T T G A A C A A C A A A T G C G T T A A C T T T A A G A C A C A G A A A A T A G C A A A G C T C T  
T T T C A G A C T C C C G C C G T C A A C T A G A A T T T A G T T T T A C T A A A A T T A G A T T T  
6601 C A A T A T T T A A A A C A C T T G T G C G A T T T T G A A A T A G A A T A T C A T A T T A C A A T G  
6661 C C A G T A A A A A A T T T T A G C A A T T G A T T T A C T G A A A C G A A A A T T T G A T T T T C  
6721 A A C G A A A T T T C A A T T T T A A A A A A A A A G C A A A A T A C A T C A A C A T C A A  
6781 G A T T T A A G C A A A T T G T A A C T A T T A A A A A T A A T C A A A A T T G T C C A C A G A T T  
6841

Repeat 5

6901 A C A C G G T T G C C G C G C A A A G T T C A A A A C A G C G T C C A C C T C T T T T C T G C G A C T T  
G A A A A C A A G G A T C G G T T A G A T T T T C C C C A A A A T T A A A T T A A A T T C A G A T C A T C  
7021 C G C G T C T C A A A A A G T T C T A C T T T T G C C G A A A C A C A A A A T C G C C G T C T A A A A T C  
R L L E K F Y F M F A R M E A V  
490 500

7081 A C T G T G A T C A T T C C C C A A T T T A A T T G T C T A T C T T C C C C C A A G C T C T T T C G C  
7141 C A A T A T T A T T A A A C A C T T G T G C G A T T T A T C C T A T A C T A C A T T C A C T T T A T C A T T  
7201 A T C A T T C T C T C C C A T T T C A C A C A T T C C A T T C T C A C A G A T C A T A A A A T T A G A C  
7261 G T T G T G T C T C A A A C C A T A A A T T T A A A A A T T T T T G A A A T T G A T A G T A T T G T G  
7321 G C C C A C A T A T A T G T A T G T A C T A C T C A C A C A A A T A A T T C A A A A T C A T C A C C  
7381 C C A A C C C A A C C A C T A C C G T A C A A T A C T T T T G C C G A A A A T A C A T T G A A T A A T  
7441 T T A C C T A T T T T G T C C C A C A A A A A T C T A A T T T G A A A T T A A C A A T A G A A T T C C A C  
7501 T C C C C G T C G A A T G C C C C G C C C T T A A A G T T G A A C A A T T G C A A A T T A G A T A T  
7561 A A A T T T G T A G T C C C C C C A T C A T T T G C C C C C A C A T C T C A A A T T G C A T T C T T T T T C  
7621 C C T G A T A C C G A T T C T G C C A G A A G A T C

# B

Repeat 1

```

ced-3 (1a, for)      1       10       20       30       40       50       60       70
GTATTAAAGGAATCAAAATTTCTGTAAGATGGCTACTGCGCAACATATTTCGACG-CAAAAATTTCTGCTAGCG
ced-3 (1a, rev)      -----AAAATTCAGAAATGGCTATTACAGTC-ATAATTTGGCGCGCAAAAATATTCTGCTAGCT
ced-3 (1b, for)      -----AAATTCGTAAGATGGCGATTACTCAACATATTTCGACGCG-AAAATTTCTGCTAGCG
ced-3 (1b, rev)      -----
fem-1 (for)          GTATTACGGCAAGAATATTATTCAGATGGCTATTTCGCAACCATATGTTTCGACGCGCAAAAATATTCTGCTAGCG
fem-1 (rev)          GTATACCGGTAAACAACAATTTCTGTAAGATGGCTATTTCGCAACCATATGTTTCGACGCGCAAAAATATTCTGCTAGCG
hlh-1 (for)          GTATTACGGGATGACAAAATTTCTGAAATGGCTACTGCGCAACATATTTCGCGCAAAAATATTCTGCTATG
hlh-1 (rev)          -----GGGAGCGCAAAAATTTCTGATATAGAAAT-CCGTATAA-----GCACAAAATATTTCGCTAGCG

consensus            -TAT-A-GG-A-A-A-AATT--GA-ATG-----A-T-C-TG-----A-A-TTG-CG-CAAAAAT-T-T-G-A-C-

                        80       90       100       110       120       130       140
ced-3 (1a, for)      AAAAATCAAGATGATCTTTTAAATGACTACTGTAGCGCG-----TTGTGTGCA-TTTCAGGGCTCAAT-
AGACTACAGATATCTTTTAAATGACTACTGTAGCG-----TTGTGCA-TTTCAGGGCTCAATTTGCAAA
ced-3 (1a, rev)      AAAA-TACAGTACACCTTTTAAATGACTACTGTAGCG-----TTGTGCA-TTTCAGGGC--TGATTTTGCAGAA
ced-3 (1b, for)      AAAACTACTGTACTTCTTTAAAGAGTACTGTAGCG-----TTGTGCA-TTTCAGGGC--TGATTTTGCAGAA
ced-3 (1b, rev)      AAAACTACAGTAACTCTTTGAGACTACTGTAGCG-----TTGTGTGCA-TTTCAGGGCTCAATTT
fem-1 (for)          AAAACTCAAGTATGTTGCTGTAATGAATAGCTAGGGTGC-----TTTGTGA-TTTCAGGGC-TG
fem-1 (rev)          AAAACTCAAGATATCTGCTGTAATGAATAGCTAGGGTGC-----TTTGTGA-TTTCAGGGC-TG
hlh-1 (for)          AAAACTCAAGATATCTGCTGTAATGAATAGCTAGGGTGC-----TTTGTGA-TTTCAGGGC-TG
hlh-1 (rev)          AAAACTCAAGATATCTGCTGTAATGAATAGCTAGGGTGC-----TTTGTGA-TTTCAGGGC-TG

consensus            A-AACTAC-GT-A-----A-G--TA-GTAG-----T-T-GT-----TTTACG-----TT-GAAA

Repeat 2
1       10       20       30       40       50       60       70       80
ced-3 (for)          TCATTCAAGATATGCTTTATTAAACATATTAATTATCATATATTCGTAATTTCTTGTAGAAATTTTGGGCTTTTCGCTCTGAT
ced-3 (rev)          TCATTCAAGATATGCTTTATTAAACATCTTAATCAATAAAGAGTAATCTTCGTAAGAAATTTTGG--TTTTCGCTCTA-

consensus            TCATTCAAGATATGCTTTATTAAAC--TATAATTATCA-TAA-G-AAT-TC-TGTG-AGAAATTTTGG--TTTTCG-TCCTA-

                        90       100       110       120       130       140       150       160
ced-3 (for)          TATGCTCTACTTTTGAATTTGCTCAACGAAAAAAT-----CATGTTGTTTGTTCATGAAATGCAAAAAAT
ced-3 (rev)          TATTTCTCTACTTTTAAAGTTGCTCAACGAAAAAATGAGGGTTAATCATGTG-----ATGTTGAAAAATA

consensus            TAT-CTCTACTTTT-AA-TTGTCTCAACGAAAAAAT-----CATGTG-----ATG--GAAAAATA

                        170       180       190       200       210       220       230       240
ced-3 (for)          GCAAA-----T-T-TTTTATATATTTT-CCGCTATTCATGTTGTGCGAGAAAAATGTAAAAAAGCGCATGCAATTTTT-
ced-3 (rev)          CAAAAAATGTATTTTAAATCAATTTTCCGCTATTCAT-TTGTGCGAGAAAGT-GTAAAAAAGCGCATGCAATTTTTTACAT

consensus            --AA-----T-T-TTTT-ATA-ATTTTT-CCGCTATTTCAT-TTGTGCGAGAAAA-T-GTAAAAA-CCGATGCAATTTTTT-

                        250       260       270       280       290
ced-3 (for)          -----CGACA-TTTTTTACATGTCGGAAGCAGACGCTTCACATGCTGGAAGCAGAGAGCG
ced-3 (rev)          TATTGACATTTTTTTACATGTCGAGCAGATGCCATTTTCACATGCTGGAAGCAGAGAGCG

consensus            -----CGACA-TTTTTTACATGTCGGA-A-C-C-CA-TTCACATGCTGGAAGCAGAGAGCG

Repeat 3
1       10       20       30       40       50       60
ced-3 (for)          CAGCTTCGAGAGTTTGG-AAAATACAGTACTCTCTTAAAGGGCGACACCCATTTGCATTGG
ced-3 (rev)          CAGCAACAATATTTTGG-AAAATACAGTAACTTTTAAAGGGCGACACCC-----
11n-12 (for)          CTCGACAGAAAGTTG-AAAATACAGTACTCTCTTAAAGGGCGACAT-
11n-12 (rev)          -----GTTAG-AAAATAGAGTACTCTTTAAAGGGCGACAC-
B0303 (1)             CAGCAACAATAAGTTTG-AAAATACAGTCTCTTTTAAAGCGACACAC-TTTTTACATT-T
B0303 (2)             CAGAAGCGAAAAATTTG-AAAATACAGTACTCTTTTAAAGGGCGACA-CCCCGCTTATATCTA
ZK643 (1)              GATTGTTT-GAAAAATTACAGTAACTTTTAAAGGGCGACAC-CGTTTGATTTAT-
ZK643 (2, for)        -----
ZK643 (2, rev)        -----TTTTTAACTACAGTACTCTTTTAAAGGGCGACA-TTTTTTGTGATTT
ZK643 (3)             -----ACCGTTTG-AAAATACAGTACTCTTTTAAAGGGCG-
gfp-1 (for)           -----
gfp-1 (rev)           -----

consensus            C-G-C--C-----T-----AAA-TA-AGT-----TTA--GC-A-CA--TTT-----

                        70       80       90       100       110
ced-3 (for)          ACCAAAAAATTTGCTGTGTCGAGACGAGTA-CCGATGTTTGTG-TC-----GCAAAA
ced-3 (rev)          -----ACAAA-TTGTGCTGTGCGAGACGGGGG-CCACA-----
11n-12 (for)          ACAAACAAATTTGCTGTGTCGAGACGGGTA-CCGATGTTTAAATTT-----GCAAA
11n-12 (rev)          -----
B0303 (1)            ATCGAAAAATTTGCTGTGTCGAGACGGGTAG-TAAATTTATGCG-CAAAAA-
B0303 (2)            ACGAAAAAATGCTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (1)             ATGAGAG-TTGTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (2, for)        ACGAAAAA-TTCTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (3)             -----TTTGTGCTGTGTCGAGACGGTGA-
gfp-1 (for)          ACAAATTTTGTGTCGCGGAGACCTGATTA-CCGATGTTTGTGCTCAAGATTACTAGG
gfp-1 (rev)          -----GTTTGTGCT-

consensus            A-----T-TGCG--CGAGAC-----C-TTT-----A-A-----

Repeat 4
1       10       20       30       40       50       60       70
ced-3 (for)          ACCAACTACGATCTGTGCGATCTCGCGCCACTTCATCGATTTGGTTTGAAGTGGGCGAGGTGAATTTGCTATTTGGTC
ced-3 (rev)          ACCAACTATGGAGCTTCGGAATTTGCTACTTAACTGATTTGGTTGAAGAAATGGCGAGCGGAATTTGCTGATTTGGC
11n-12 (for)          ACAAATATGCTGTGCTGTGCGAGACGGGTA-CCGATGTTTAAATTT-----GCAAA
11n-12 (rev)          -----
B0303 (1)            ATCGAAAAATTTGCTGTGTCGAGACGGGTAG-TAAATTTATGCG-CAAAAA-
B0303 (2)            ACGAAAAAATGCTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (1)             ATGAGAG-TTGTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (2, for)        ACGAAAAA-TTCTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (3)             -----TTTGTGCTGTGTCGAGACGGTGA-
gfp-1 (for)          ACAAATTTTGTGTCGCGGAGACCTGATTA-CCGATGTTTGTGCTCAAGATTACTAGG
gfp-1 (rev)          -----GTTTGTGCT-

consensus            A-----T-TGCG--CGAGAC-----C-TTT-----A-A-----

Repeat 5
1       10       20       30       40       50       60
ced-3 (for)          TTTTAAAG-CACACGAAAAAATAGCGAGAGGCTCTTTTCGAAAGTGGGCGAGGTGAATTTGCTATTTGGTC
ced-3 (rev)          TTTTAAAGCGCAGACGAAAAAGAGGCGAGGCTCTTTTCGAAATTTGCTGCGGGCGCAAC
11n-12 (for)          ACAAATATGCTGTGCTGTGTCGAGACGGGTA-CCGATGTTTAAATTT-----GCAAA
11n-12 (rev)          -----
B0303 (1)            ATCGAAAAATTTGCTGTGTCGAGACGGGTAG-TAAATTTATGCG-CAAAAA-
B0303 (2)            ACGAAAAAATGCTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (1)             ATGAGAG-TTGTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (2, for)        ACGAAAAA-TTCTGCTGTGTCGAGACGGGTA-CCGATGTTTGTGCGCAAAAATTTGCTAT
ZK643 (3)             -----TTTGTGCTGTGTCGAGACGGTGA-
gfp-1 (for)          ACAAATTTTGTGTCGCGGAGACCTGATTA-CCGATGTTTGTGCTCAAGATTACTAGG
gfp-1 (rev)          -----GTTTGTGCT-

consensus            TTT-AAG-----CACGAAAAA-AGCG-GAG-C-T-TTTTGA-C-G-TGCGGCGG-CAAC

```

Figure 4. *ced-3* Genomic Organization, Nucleotide Sequence, and Deduced Amino Acid Sequence

(A) The genomic sequence of the *ced-3* region, as obtained from plasmid pJ107, which rescued the *Ced-3* mutant phenotype. The deduced amino acid sequence of the CED-3 protein is derived from the DNA sequence of the *ced-3* cDNA clone pJ87 and from other experiments described in the text and in Experimental Procedures. The likely start site of translation is marked with a closed arrowhead. The 5' end of pJ87 is indicated by an open arrowhead. The SL1 splice acceptor site of the *ced-3* transcript is boxed. The *S*all site is represented by a bracket. The positions of 12 *ced-3* mutations are indicated. Repetitive elements are indicated as arrows above the relevant sequences. Numbers on the left indicate nucleotide positions, beginning with the start of pJ107. Numbers below the amino acid sequence indicate amino acid positions.

(B) Comparison of repetitive elements in *ced-3* with repetitive elements in the genes *fem-1*, *hlh-1*, *lin-12*, and *gip-1* and the cosmid B0303 and ZK643 (see text for references). In the case of inverted repeats, each arm of a repeat (for forward or reverse, respectively) was compared with both its partner and with individual arms of the other repeats. Dashes represent gaps in the sequence to allow optimal alignment, except in the consensus sequence, in which dashes indicate nonidentical residues.

untranslated sequence, not all of which is essential for *ced-3* function, since genomic constructs that lack the last 380 bp of the 3'-most region (pJ107 and its derivatives; see Figure 1B) rescued the *Ced-3* mutant phenotype. The pJ87 cDNA clone ends with a poly(A) sequence (data not shown), suggesting that the 3' end of pJ87 corresponds to the extreme 3' end of the *ced-3* transcript. The 5' end of pJ87 does not contain trans-spliced sequences (Bektesh et al., 1988; Huang and Hirsh, 1989) and therefore might or might not include the 5' end of the *ced-3* transcript.

To determine the 5' end of the *ced-3* transcript, we performed primer extension experiments (Figure 5). Two primers containing sequences separated by 177 bp in the

genomic DNA sequence (see below) and by 123 bp in the cDNA sequence were used for the primer extension reactions: Pex1, starting at position 2305 of the genomic sequence, and Pex2, starting at position 2482 of the genomic sequence. The Pex2 reaction yielded two major products of 283 nt and 409 nt, whereas the Pex1 reaction gave one product of 160 nt. The 160 nt product of the Pex1 reaction corresponds to the 283 nt product of the Pex2 reaction, since these products differ in size by 123 nt. Products of these lengths are consistent with the presence of a *ced-3* transcript that is trans-spliced to a 22 bp *C. elegans* spliced leader (Bektesh et al., 1988) at a consensus splice acceptor site at position 2166 of the genomic

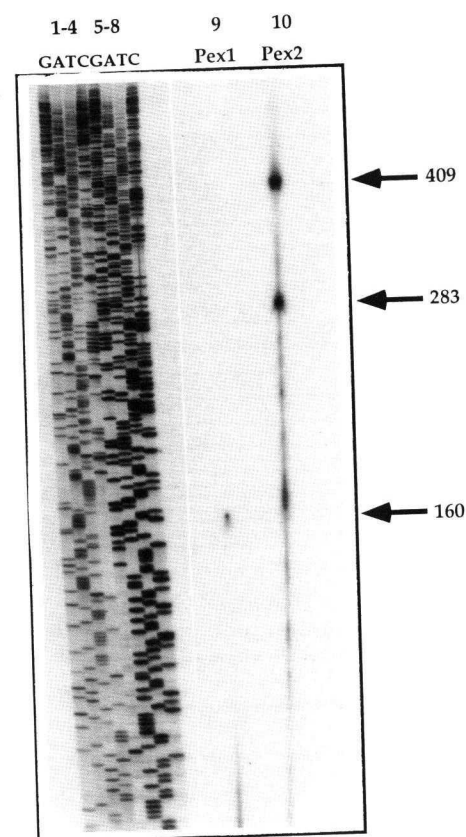


Figure 5. A Primer Extension Experiment to Determine the *ced-3* Transcription Initiation Site

See text for details. Lanes 1–4, DNA sequencing reaction products using primer Pex1 and pJ40 as the template. Lanes 5–8, sequencing reaction products using the primer Pex2 and pJ40 as the template. Lane 9, primer extension reaction products using primer Pex1 and N2 total RNA as template. Lane 10, primer extension reaction products using primer Pex2 and N2 total RNA as template. Sizes of the primer extension products are indicated to the right of lane 10.

sequence. The larger Pex2 product might be a result of priming by the Pex2 primer from an mRNA other than the *ced-3* mRNA or might identify a second *ced-3* transcript. We failed to amplify a *ced-3* transcript using primers located between positions 1 and 2166 of the genomic sequence (see Figure 4A) and primers located in the cDNA (data not shown), indicating that if an alternate *ced-3* transcript exists, it is not entirely encoded by DNA within the minimal rescuing fragment.

If a trans-spliced *ced-3* mRNA exists, it should be possible to use the polymerase chain reaction to amplify a *ced-3* product from total *C. elegans* RNA using primers specific for spliced leader sequences and *ced-3* cDNA sequences. We therefore prepared cDNA from total *C. elegans* RNA using reverse transcriptase and amplified the cDNAs using the primers SL1 (which contains the SL1 spliced leader sequence; Bektesh et al., 1988) and log-5 (which starts at position 2897 of the genomic sequence). The products of this reaction were reamplified using the primers SL1 and oligo10 (which starts at position 2344 of the genomic sequence). A product of the expected length (~200 bp)

Table 1. Sequences of *ced-3* Mutations

Allele	Wild-Type Sequence	Mutant Sequence	Substitution or Splice Site Change
n717	ttttgcag/CAA	ttttgca/CAA	Exon 7 acceptor
n718	GGA	AGA	G65R
n1040	CTC	TTC	L27F
n1129, n1164	GCA	GTA	A449V
n1163	TCC	TTC	S486F
n1165	CAG	TAG	Q404amber
n1286	TGG	TGA	W428opal
n1949	CAA	TAA	Q412ochre
n2426	GAG	AAG	E483K
n2430	GCT	GTT	A466V
n2433	GGC	AGC	G360S

Amino acid positions correspond to the numbering in Figure 4A.

was cloned into the PCR1000 vector (Invitrogen, San Diego, California), and its sequence was determined (data not shown). This sequence confirmed that at least some *ced-3* transcripts are trans-spliced to the *C. elegans* spliced leader SL1 at a consensus splice acceptor at position 2166 of the genomic sequence. Similar experiments using a primer containing the SL2 spliced leader sequence (Huang and Hirsh, 1989) failed to identify an SL2 trans-spliced mRNA. Since the primer extension experiments identified a major *ced-3* mRNA that probably is trans-spliced, it seems likely that a large proportion of *ced-3* mRNA is SL1 trans-spliced. Based upon these observations, we propose that the translational start codon of *ced-3* is the first in-frame ATG downstream of the SL1 splice acceptor site at position 2232 of the genomic sequence and that the CED-3 protein is 503 amino acids in length (see Figure 4A).

To define the structure of the *ced-3* gene, we determined the genomic sequence of the *ced-3* gene from the plasmid pJ107. The insert in pJ107 is 7653 bp in length (see Figure 4A). Comparison of the *ced-3* genomic and cDNA sequences revealed that the *ced-3* gene has seven introns that range in size from 54 bp to 1195 bp. The four largest introns, as well as sequences 5' of the start codon (see below), contain repetitive elements, some of which have been previously identified in noncoding regions of other *C. elegans* genes, such as *fem-1* (Spence et al., 1990), *lin-12* (Yochem and Greenwald, 1989; J. Yochem and Greenwald, personal communication), *glp-1* (Yochem and Greenwald, 1989), and *hlf-1* (Krause et al., 1990), as well as in the cosmid ZK643 and B0303 (Sulston et al., 1992) (see Figure 4B). Genomic sequence analysis of a *ced-3* homolog from the related nematode *C. briggsae* (J. Y. and S. L., unpublished data) revealed that these repeats are not present in this nematode species, suggesting that the repeats do not have a role in regulating *ced-3* expression. It is possible that such repeats represent active or inactive transposable elements.

The predicted CED-3 protein is hydrophilic (256 of 503 residues are charged or polar) and does not contain any obvious potential transmembrane domains. One region of

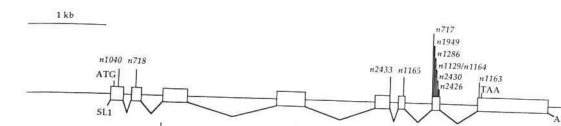


Figure 6. Intron-Exon Structure of the *ced-3* Gene

Intron-exon structure (represented by lines and boxes, respectively) of the *ced-3* gene and positions of 12 *ced-3* mutations. The trans-spliced SL1 leader, the serine-rich region, the mutations, and the presumptive translational start site (ATG), termination site (TAA), and polyadenylation site (AAA) are indicated.

the CED-3 protein is rich in serines: from amino acid 107 to amino acid 205, 32 of 99 amino acids are serines.

We determined the sequences of 12 EMS-induced *ced-3* mutations (see Figure 4A; Table 1). Eight are missense mutations, three are nonsense mutations, and one alters a conserved G at the presumptive splice acceptor site of intron 6.

To identify functionally important regions of the CED-3 protein, we cloned and determined the genomic sequences of the *ced-3* genes from the related nematode species *C. briggsae* and *C. vulgaris*. Sequence comparisons showed that amino acids corresponding to residues 1–205 of the *C. elegans* CED-3 protein are less conserved

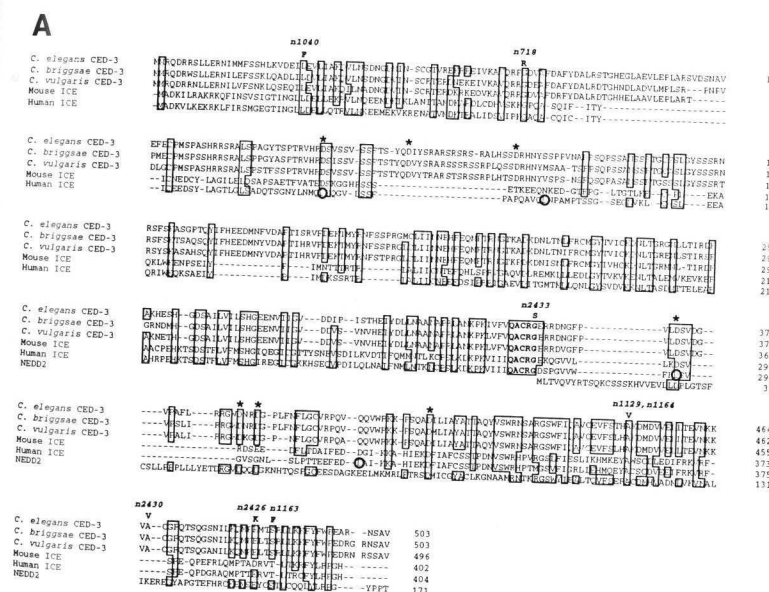


Figure 7. CED-3 Protein

(A) Comparison of the CED-3 protein sequences from *C. elegans*, *C. briggsae*, and *C. vulgaris* with the human and mouse ICEs and with the mouse Nedd-2 protein. Amino acids are numbered to the right of each sequence. Dashes indicate gaps in the sequence to allow optimal alignment. Residues that are identical among more than half of the proteins and between nematode and mammalian sequences are boxed. Missense *ced-3* mutations are indicated above the comparison blocks, showing the residue in the mutant CED-3 protein and the allele name. Asterisks indicate potential aspartate self-cleavage sites in the CED-3 protein. Circles indicate known aspartate self-cleavage sites in human ICE. Residues indicated in boldface correspond to the highly conserved pentapeptide containing the active cysteine in ICE.

(B) Comparison of structural features of the CED-3 protein and human ICE. The predicted proteins corresponding to the ICE proenzyme and CED-3 are represented. The active site in ICE and the predicted active site in CED-3 are indicated by the closed rectangles. The four known cleavage sites in ICE flanking the processed ICE subunits (p24, which was detected in low quantities when ICE was purified [Thornberry et al., 1992], p20, and p10) and two conserved presumptive cleavage sites in the CED-3 protein are indicated with solid lines and linked with dotted lines. Five other potential cleavage sites in the CED-3 protein are indicated with dashed lines. The positions of the aspartate (D) residues at potential cleavage sites are indicated below each diagram. The carboxyl terminus of p24 has not been determined and is indicated by a dotted arrow.

amino acids 166–287 of the human ICE protein: 49 of 115 residues are identical (43% identity). Cys-285 is thought to be an essential component of the active site of ICE (Thornberry et al., 1992). The amino acid pentapeptide QACRG containing this active cysteine is the longest peptide conserved among the murine and human ICE proteins and the CED-3 proteins of the three nematode species.

Active human ICE is composed of two subunits (p20 and p10) that appear to be proteolytically cleaved from a single proenzyme by the mature enzyme (Thornberry et al., 1992). Four cleavage sites in the proenzyme have been defined. Only p20 and p10 are necessary for the *in vitro* enzymatic activity of ICE, suggesting that the three additional fragments resulting from ICE cleavage are not required for ICE function. Two of these cleavage sites (aspartic acid–serine dipeptides at positions 103–104 and 297–298 of ICE) are conserved in CED-3 (positions 131–132 and 371–372, respectively) (Figure 7).

The carboxy-terminal portion of the CED-3 protein and the p10 subunit of ICE are similar to the protein product of the murine gene *nedd-2* (Figure 7A), which is highly expressed during embryonic brain development and is down-regulated in adult brain (Kumar et al., 1992). The *C. elegans* CED-3 protein and the Nedd-2 protein are 27% identical, as are the ICE and Nedd-2 proteins (Figure 7A). The Nedd-2 protein apparently does not contain the QACRG peptide found at the active site of ICE (Figure 7A). Six of the eight known *ced-3* missense mutations (*n718*, *n1040*, *n1129*, *n1164*, *n2426*, and *n2433*) alter amino acids that are identical among the three nematode CED-3 proteins and human ICE. For example, the mutation *n2433* introduces a glycine to serine substitution at an absolutely conserved glycine near the putative active site cysteine (see Figure 4A; Figure 7A). Four mutations (*n1129*, *n1163*, *n1164*, and *n2426*) alter amino acids that are identical among the nematode CED-3 proteins and the Nedd-2 protein (Figure 7A).

## Discussion

The genes *ced-3* and *ced-4* are the only genes known to be required for programmed cell death to occur in *C. elegans* (Ellis and Horvitz, 1986). Our genetic and molecular studies of the *ced-3* gene have revealed that this gene shares a number of features with *ced-4*: like *ced-4* (see Yuan and Horvitz, 1992), *ced-3* is not required for viability and is expressed mostly during embryogenesis, the stage during which 113 of the 131 programmed cell deaths occur (Sulston et al., 1983). Furthermore, just as *ced-3* gene function is not required for *ced-4* gene expression (Yuan and Horvitz, 1992), *ced-4* gene function is not required for *ced-3* gene expression. Thus, these two genes do not appear to control the onset of programmed cell death by acting sequentially in a transcriptional regulatory cascade.

The CED-4 protein is novel in sequence, and the only hint concerning its function is that two regions of the protein show some similarity to the EF-hand motif, which binds calcium (Yuan and Horvitz, 1992). For this reason we have suggested that the CED-4 protein and hence programmed cell death in *C. elegans* might be regulated by

calcium. However, no direct evidence for this hypothesis has yet been obtained. The CED-3 protein similarly contains a region that offers a clue about possible function: a region of 99 amino acids contains 32 serines. Since serines are common phosphorylation sites (Edelman et al., 1987), it is possible that the CED-3 protein and hence programmed cell death in *C. elegans* are regulated by phosphorylation. Phosphorylation has previously been suggested to function in cell death (McConkey et al., 1990). Although the precise sequence of the serine-rich region varies among the three *Caenorhabditis* species studied, the relatively high number of serines is conserved (32, 31, and 33 in *C. elegans*, *C. briggsae*, and *C. vulgaris*, respectively). None of the mutations in *ced-3* affects the serine-rich region. These observations are consistent with the hypothesis that the presence of serines is more important than the precise amino acid sequence within this region.

Much more striking than the presence of the serine-rich region in the CED-3 protein is the similarity among the non-serine-rich regions of CED-3 and the human and murine ICEs. Human ICE was identified as a substrate-specific protease that cleaves the 31 kd pro-IL-1 $\beta$  between Asp-116 and Ala-117 to produce the mature 17.5 kd IL-1 $\beta$ . IL-1 $\beta$  is a cytokine involved in mediating a wide range of biological responses, including inflammation, septic shock, wound healing, hematopoiesis, and the growth of certain leukemias (Dinarello, 1991; diGiovine and Duff, 1990). A specific inhibitor of ICE, the *crmA* gene product of cowpox virus prevents the proteolytic activation of IL-1 $\beta$  (Ray et al., 1992) and inhibits the host inflammatory response (Ray et al., 1992). Cowpox virus carrying a deleted *crmA* gene is unable to suppress the inflammatory response of chick embryos, resulting in a reduction in the number of virus-infected cells and less damage to the host (Palumbo et al., 1989). These observations indicate the importance of ICE in bringing about the inflammatory response.

A region of 115 amino acids (residues 246–360 of CED-3) shows the highest identity (43%) between the *C. elegans* CED-3 protein and the human ICE protein. This region contains a conserved pentapeptide QACRG (positions 356–360 of the CED-3 protein), which contains a cysteine known to be essential for ICE function. Specific modification of this cysteine in human ICE results in a complete loss of activity (Thornberry et al., 1992). The *ced-3* mutation *n2433* alters the conserved glycine in this pentapeptide and eliminates *ced-3* function, suggesting that this glycine is important for *ced-3* activity and might be an integral part of the active site of ICE. Six of the other seven identified *ced-3* missense mutations also affect highly conserved residues that are likely to be important for the actions of both CED-3 and ICE. Interestingly, the mutations *n718* (position 65 of CED-3) and *n1040* (position 27 of CED-3) eliminate *ced-3* function *in vivo* yet alter conserved residues that are not contained in either the mature p10 or p20 subunits of ICE (Thornberry et al., 1992). It is possible that these residues have a noncatalytic role in both CED-3 and ICE function, for example, in maintaining a proper conformation for proteolytic activation. The hu-

man ICE proenzyme (p45) can be proteolytically cleaved at four sites (Asp-103, Asp-119, Asp-297, and Asp-316 of ICE) to generate two peptides (p20 and p10) necessary for *in vitro* activity (Thornberry et al., 1992) and three other peptides with as yet undefined functions. At least two of these cleavage sites are conserved in CED-3, indicating that the CED-3 protein might be processed as well.

The similarity between the CED-3 and ICE proteins strongly suggests that CED-3 functions as a cysteine protease in controlling programmed cell death by proteolytically activating or inactivating a substrate protein or proteins. A potential substrate for CED-3 might be the product of the *ced-4* gene. The CED-4 protein contains six aspartate residues that might be targets of the CED-3 protein. Four of these aspartates are followed by a serine (Asp-151, Asp-184, Asp-192, and Asp-541), and two are followed by an alanine (Asp-25 and Asp-459); of the four ICE cleavage sites in the ICE proenzyme, two are aspartic acid–serine and one is aspartic acid–alanine. Alternatively, the CED-3 protein might directly cause cell death by proteolytically cleaving proteins that are crucial for cell viability.

The similarity between CED-3 and ICE defines a novel protein family. Thornberry et al. (1992) suggested that the sequence GDSPG at position 287 of ICE resembles a GX(S/C)XG motif found in serine and cysteine protease active sites. In the three nematode CED-3 proteins, however, only the first glycine of this sequence is conserved, and in mouse ICE the S/C is missing, suggesting that the CED-3/ICE family shares little sequence similarity with known protease families.

The similarity between CED-3 and ICE suggests not only that CED-3 might function as a cysteine protease but also that ICE might function in programmed cell death in vertebrates. Consistent with this hypothesis, after murine peritoneal macrophages were stimulated with lipopolysaccharide and induced to undergo programmed cell death by exposure to extracellular ATP, mature active IL-1 $\beta$  was released into the culture supernatant; by contrast, when cells were injured by scraping, IL-1 $\beta$  was released exclusively as the inactive proenzyme (Hogquist et al., 1991). These results suggest that ICE might be activated upon induction of programmed cell death. A role for ICE in programmed cell death need not be mediated by IL-1 $\beta$  but rather could be mediated by another ICE substrate. ICE transcripts have been detected in cells that do not make IL-1 $\beta$  (Cerretti et al., 1992), suggesting that other ICE substrates might well exist. Alternatively, members of the CED-3/ICE family other than ICE might function in vertebrate programmed cell death.

The p10 subunit of ICE and the carboxy-terminal portions of the CED-3 protein are similar to the protein encoded by the murine *nedd-2* gene, which is expressed during early embryonic brain development (Kumar et al., 1992). Since the Nedd-2 protein apparently lacks the QACRG active site region and is similar to ICE, primarily in the region of the p10 subunit of ICE, *nedd-2* might function noncatalytically to regulate an ICE or ICE-like p20 subunit. Interestingly, three *ced-3* missense mutations alter residues conserved among the Nedd-2 and CED-3 proteins. *nedd-2* gene expression is high during embryonic brain

development (Kumar et al., 1992), when much programmed cell death occurs (Oppenheim, 1981). These observations suggest that Nedd-2 might function in programmed cell death.

The *C. elegans* gene *ced-9* protects cells from undergoing programmed cell death by directly or indirectly antagonizing the activities of *ced-3* and *ced-4* (Hengartner et al., 1992). The vertebrate gene *bcl-2* acts functionally similarly to *ced-9*: overexpression of *bcl-2* protects or delays the onset of apoptotic cell death in a variety of vertebrate cell types as well as in *C. elegans* (Vaux et al., 1988; Nunez et al., 1990; Garcia et al., 1992; Sentman et al., 1992; Strasser et al., 1991; Vaux et al., 1992; M. Hengartner and H. R. H., unpublished data). Thus, if ICE or another CED-3/ICE family member is involved in vertebrate programmed cell death, an intriguing possibility is that *bcl-2* could act by modulating its activity. Furthermore, since *bcl-2* is a dominant oncogene (overexpression of *bcl-2* as a result of chromosomal translocation occurs in 85% of follicular and 20% of diffuse B cell lymphomas; Fukuhara et al., 1979; Levine et al., 1985; Yunis et al., 1987), ICE and other *ced-3*/ICE family members might be recessive oncogenes: the elimination of such cell death genes could prevent normal cell death and promote malignancy, just as overexpression of *bcl-2* does.

## Experimental Procedures

### General Methods and Strains

The techniques used for culturing *C. elegans* were as described by Brenner (1974). All strains were grown at 20°C. The wild-type strains were *C. elegans* variety Bristol strain N2, Bergerac strain EM1002 (Emmons et al., 1983), *C. briggsae*, and *C. vulgaris* (V. Ambros, personal communication). Genetic nomenclature follows the standard *C. elegans* system (Horvitz et al., 1979). The mutations used have been described by Brenner (1974) and by Hodgkin et al. (1988) or were isolated by us. These mutations are listed below.

LG I: *ced-1(e1735)*.

LG IV: *unc-31(e928)*, *unc-30(e191)*, *ced-3(n717)*, *n718*, *n1040*, *n1129*, *n1163*, *n1164*, *n1165*, *n1286*, *n1949*, *n2426*, *n2430*, *n2433*, *unc-26(e205)*.

LG V: *egl-1(n487)*, *n986*.

LG X: *dpy-3(e27)*.

### Isolation of Additional Alleles of *ced-3*

A noncomplementation screen was designed to isolate new alleles of *ced-3*. Because animals carrying *ced-3(n717)* in trans to a deficiency are viable (Ellis and Horvitz, 1986), we expected animals carrying a complete loss-of-function *ced-3* allele generated by mutagenesis to be viable in trans to *ced-3(n717)*, even if the new allele caused inviability in homozygotes. We used an *egl-1* mutation in our screen. Dominant mutations in *egl-1* cause the two hermaphrodite-specific neurons known as the HSNs to undergo programmed cell death (Trent et al., 1983). The HSNs are required for normal egg-laying, and *egl-1* hermaphrodites, which lack HSNs, are egg-laying defective. The mutant phenotype of *egl-1* is suppressed in a *ced-3*; *egl-1* strain because mutations in *ced-3* block programmed cell deaths. *egl-1(n986)* males were mutagenized with EMS (20 mM for 4 hr; Sigma, St. Louis, Missouri) and crossed with *ced-3(n717)* *unc-26(e205)*; *egl-1(n487)*; *dpy-3(e27)* hermaphrodites. Most cross progeny hermaphrodites were egg-laying defective because they were heterozygous for *ced-3* and homozygous for *egl-1*. Rare egg-laying competent animals were picked as candidates for carrying new alleles of *ced-3*. Four such animals were isolated from about 10,000 hermaphrodite F1 cross progeny of EMS-mutagenized animals. These animals could have carried either a dominant suppressor of the egg-laying defect of *egl-1* or a recessive mutation in *ced-3*. To distinguish between the two possibi-

## Cosmid Libraries

### Identification and Mapping of RFLPs

By probing DNA from these strains with the Tc1 insert of plasmid pCec2001 (Emmons et al., 1983), we identified a 5.1 kb EcoRI Tc1-containing restriction fragment specific to the Bristol strain and closely linked to *ced-3*. We named this RFLP *nP35*. *nP35* was localized to cosmids MMM-C1 and MMM-C9, which were used to identify a cosmid contig in the *ced-3* region (see text for details).

mid T10H5 and B0564, respectively.

We mapped the RFLPs *nP33*, *nP34*, *nP35*, *nP36*, and *nP37* with respect to the genes *unc-30*, *ced-3*, and *unc-26*. The location of *unc-30* on the physical map was determined by R. Hoskins (personal communication). Southern blots using DNA from the *Ced-3* non-*Unc-26* and *Unc-26* non-*Ced-3* inbred recombinant strains described above were used to map *nP33*, *nP34*, and *nP35* (data not shown). Three of three *Unc-26* non-*Ced-3* recombinants carried the Bristol RFLP *nP33*, while 2 of 2 *Ced-3* non-*Unc-26* recombinants showed the Bergerac pattern; thus, *nP33* maps very close to or to the right of *unc-26*. For *nP34*, 2 of 2 *Ced-3* non-*Unc-26* recombinants and 2 of 3 *Unc-26* non-*Ced-3* recombinants showed the Bergerac pattern; the remaining *Unc-26* non-*Ced-3* recombinant showed the Bristol pattern. The genetic distance between *ced-3* and *unc-26* is about 0.3 map units (mu). Thus *nP34* maps between *ced-3* and *unc-26*, about 0.2 mu to the right of *ced-3*. Similar experiments mapped *nP35*, defined by the 5.1 kb restriction fragment containing a Bristol-specific Tc1 element, to about 0.2 mu to the right of *ced-3*. Southern blots of the *Unc-30* non-*Ced-3* and *Ced-3* non-*Unc-30* inbred strains described above were used to map *nP36* and *nP37* (data not shown). *nP36* mapped very close to or to the left of *unc-30*, since 2 of 2 *Unc-30* non-*Ced-3* recombinants showed the Bristol pattern on Southern blots, and 2 of 2 *Ced-3* non-*Unc-30* recombinants showed the Bergerac pattern. Similarly, *nP37* mapped very close to or to the left of *unc-30* since 4 of 4 *Ced-3* non-*Unc-30* recombinants showed the Bergerac pattern on Southern blots, and 6 of 6 *Unc-30* non-*Ced-3* recombinants showed the Bristol pattern.

### Germline Transformation

Miniprep DNA was used when cosmid subclones were injected. Miniprep DNA was prepared from 1.5 ml of an overnight bacterial culture grown in superbroth (Sambrook et al., 1989). DNA was extracted by the alkaline lysis method, as described by Sambrook et al. (1989), and was treated with RNAase A (37°C for 30 min) and then with proteinase K (55°C for 30 min), phenol extracted and then chloroform extracted, precipitated twice (first in 0.3 M sodium acetate [pH 5.2] and then in 0.1 M potassium acetate [pH 7.2]), and resuspended in 5 µl of injection buffer as described by Fire (1986). The DNA concentrations used in injections were between 100 µg/ml and 1 mg/ml.

pJ7.4 were generated by ExoIII deletion of pJ10.7. All transformation experiments used a *ced-1*(e1735); *unc-31*(e928) *ced-3*(n717) strain. The *ced-1* mutation was present to facilitate scoring of the *Ced-3* phenotype. Mutations in *ced-1* block the engulfment of corpses generated by programmed cell death, causing the corpses of dead cells to persist much longer than in the wild type (Hedgecock et al., 1983). Thus, the presence of a corpse indicates a cell that has undergone programmed cell death. The *Ced-3* phenotype was scored by counting the number of corpses present in the head of young L1 animals. Cosmid C48D1 or plasmid subclones were mixed with C14G10 (which contains the *unc-31*(+) gene) at a weight ratio of 2:1 or 3:1 to increase the chances that an *unc-31*(+) transformant would contain the cosmid or plasmid being tested. Usually, 20–30 animals were injected in one experiment. Non-Unc F1 progeny of injected animals were isolated 3–4 days later. About one half to one third of the non-Unc progeny transmitted the non-Unc phenotype to the F2 generation and could be used to establish a line of transgenic animals. The young L1 progeny of such non-Unc transgenics were checked for the number of dead cells present in the head using Nomarski optics, as described by Ellis and Horvitz (1986). Animals of the transgenic line *nIs1* transmitted the non-Unc phenotype to 100% of their progeny, implying that the *ced-3* and *unc-31* transgenes had integrated into the genome. Non-Unc animals of the transgenic line *nEx2* transmitted the non-Unc phenotype to 50% of their progeny, implying that the *ced-3* and *unc-31* transgenes were present on an extrachromosomal array that is maintained as an unstable free duplication (Way and Chalfie, 1988).

### Isolation of *ced-3* cDNAs

### Construction of a *ced-3* Frameshift Mutation

#### Determination of the *ced-3* Transcription Initiation Site

### DNA Sequence Analysis

The CED-3 protein sequence was compared with protein sequences in the GenBank, PIR, and SWISS-PROT data bases at the National Center for Biotechnology Information using the BLAST network service.

### Cloning of *ced-3* Genes from Other Nematode Species

The *C. briggsae* and *C. vulgaris* *ced-3* genes were isolated from corresponding phage genomic libraries (C. Link, personal communication) using the *ced-3* cDNA subclone pJ118 as a probe under low stringency hybridization conditions (5 × SSPE [Sambrook et al., 1989], 20% formamide, 0.02% Ficoll, 0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone, 1% sodium dodecyl sulfate at 40°C overnight) and washed in 1 × SSPE and 0.5% sodium dodecyl sulfate twice at room temperature and twice at 42°C for 20 min for each wash.

## Acknowledgments

J. Y. and S. S. contributed equally to this paper. Correspondence should be addressed to H. R. H. We thank John Sulston and Alan Coulson for mapping our *ced-3* cosmids; Chris Link for *C. briggsae* and *C. vulgaris* genomic libraries; Chad Nusbaum for providing a filter with stage-specific RNAs; Roger Hoskins for providing us with the location of *unc-30* on the physical map and for our use of the cosmid C14G10; Michael Hengartner for isolating the mutations *n2426*, *n2430*, and *n2433*; Ron Ellis for isolating the mutation *n1949*; Erika Hartwig and Erik Jorgensen for help with Figures 2, 4, and 5; and Andrew Chisholm, Michael Basson, Michael Hengartner, and Yishi Jin for helpful comments about the manuscript. H. R. H. and the research performed at the Massachusetts Institute of Technology were supported by National Institute of Health (NIH) research grant GM24663 from the United States Public Health Service and by the Howard Hughes Medical Institute (HHMI). The research performed at the Massachusetts General Hospital was supported by the Cardiovascular Research Center of the Massachusetts General Hospital. J. Y. was supported by a scholarship from Harvard University, as a research associate of HHMI, and by the Cardiovascular Research Center of the Massachusetts General Hospital. S. S. was supported by a William Keck Foundation fellowship and by an NIH training grant. S. L. was a research fellow of the Medical Research Council of Canada. H. R. H. is an Investigator of HHMI.

Received July 16, 1993; revised September 6, 1993.

## References

- Austin, J., and Kimble, J. (1987). *g/p-1* is required in the germline for regulation of the decision between mitosis and meiosis in *C. elegans*. *Cell* 51, 589-599.
- Bektesh, S., Van Doren, K., and Hirsh, D. (1988). Presence of the *Caenorhabditis elegans* spliced leader on different mRNAs and in different genera of nematodes. *Genes Dev.* 2, 1277-1283.
- Black, R. A., Kronheim, S. R., and Sleath, P. R. (1989). Activation of interleukin-1 $\beta$  by a co-induced protease. *FEBS Lett.* 247, 386-390.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71-94.
- Cerretti, D. P., Kozlosky, C. J., Mosley, B., Nelson, N., Ness, K. V., Hogenkamp, C. O., and Smith, W. L. (1990). The structure of interleukin-1 $\beta$  determined from the complementary DNA sequence. *J. Biol. Chem.* 265, 1155-1160.
- Hongo, K., Nett, M. A., Unanue, E. R., and Chaplin, D. D. (1991). Interleukin 1 is processed and released during apoptosis. *Proc. Natl. Acad. Sci. USA* 88, 8485-8489.
- Horvitz, H. R., Brenner, S., Hodgkin, J., and Herman, R. K. (1979). A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. *Mol. Gen. Genet.* 175, 129-133.
- Huang, X. Y., and Hirsh, D. (1989). A second *trans*-spliced RNA leader sequence in the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* 86, 8640-8644.
- Ish-Horowitz, D., and Burke, J. F. (1981). Rapid and efficient cosmid cloning. *Nucl. Acids Res.* 9, 2989-2998.
- Kim, S. K., and Horvitz, H. R. (1990). The *Caenorhabditis elegans* *g/p-1* gene is required for the germline decision between mitosis and meiosis. *Cell* 61, 1311-1320.

- gene *lin-10* is broadly expressed while required specifically for the determination of vulval cell fates. *Genes Dev.* 4, 357-371.
- Kostura, M. J., Tocci, M. J., Limjuco, G., Chin, J., Cameron, P., Hillman, A. G., Chartrain, N. A., and Schmidt, J. A. (1989). Identification of a monocyte specific pre-interleukin-1 $\beta$  convertase activity. *Proc. Natl. Acad. Sci. USA* 86, 5227-5231.
- Krause, M., and Hirsh, D. (1984). Actin expression in *Caenorhabditis elegans*. In *Molecular Biology of the Cytoskeleton*, G. G. Borisy, D. W. Cleveland, and D. B. Murphy, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 287-292.
- Krause, M., Fire, A., Harrison, S. W., Priess, J., and Weintraub, H. (1990). CeMyoD accumulation defines the body wall muscle cell fate during *C. elegans* embryogenesis. *Cell* 63, 907-919.
- Kumar, S., Tomooka, Y., and Noda, M. (1992). Identification of a set of genes with developmentally down-regulated expression in the mouse brain. *Biochem. Biophys. Res. Commun.* 185, 1155-1161.
- Levine, E. G., Arthur, D. C., Frizzera, G., Peterson, B. A., Hurd, D. D., and Bloomfield, C. D. (1985). There are differences in cytogenetic abnormalities among histologic subtypes of the non-Hodgkin's lymphomas. *Blood* 66, 1414-1422.
- Martin, D. P., Schmidt, R. E., Distefano, P. S., Lowry, O. H., Carter, J. G., and Johnson, E. M., Jr. (1988). Inhibitors of protein synthesis and RNA synthesis prevent neuronal death caused by nerve growth factor deprivation. *J. Cell Biol.* 106, 829-844.
- McConkey, D. J., Orrenius, S., and Jondal, M. (1990). Agents that elevate cAMP stimulate DNA fragmentation in thymocytes. *J. Immunol.* 145, 1227-1230.
- Meneely, P. M., and Herman, R. K. (1979). Lethals, steriles and deficiencies in a region of the X chromosome of *Caenorhabditis elegans*. *Genetics* 92, 99-105.
- Nunez, G., London, L., Hockenbery, D., Alexander, M., McKearn, J. P., and Korsmeyer, S. J. (1990). Deregulated *Bcl-2* gene expression selectively prolongs survival of growth factor-deprived hematopoietic cells lines. *J. Immunol.* 144, 3602-3610.
- Oppenheim, R. W., (1981). Neuronal cell death and some related regressive phenomena during neurogenesis: a selective historical review and progress report. In *Studies in Developmental Neurobiology*, W. M. Cowan, ed. (New York: Oxford University Press), pp. 74-133.
- Palumbo, G. J., Pickup, D. J., Fredrickson, T. N., McIntyre, L. J., and Buller, R. M. (1989). Inhibition of an inflammatory response is mediated by a 38-kDa protein of cowpox virus. *Virology* 171, 262-273.
- Ray, C. A., Black, R. A., Kronheim, S. R., Greenstreet, T. A., Sleath, P. R., Salvesen, G. S., and Pickup, D. J. (1992). Viral inhibition of inflammation: cowpox virus encodes an inhibitor of the interleukin-1 $\beta$  converting enzyme. *Cell* 69, 597-604.
- Ruvkun, G., Ambros, V., Coulson, A., Waterston, R., Sulston, J., and Horvitz, H. R. (1988). Molecular genetics of the *Caenorhabditis elegans* heterochronic gene *lin-14*. *Genetics* 121, 501-516.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).
- Sentman, C. L., Shutter, J. R., Hockenbery, D., Kanagawa, O., and Korsmeyer, S. J. (1992). *bcl-2* inhibits multiple forms of apoptosis but not negative selection in thymocytes. *Cell* 67, 879-888.
- Spence, A. M., Coulson, A., and Hodgkin, J. (1990). The product of *fem-1*, a nematode sex-determining gene, contains a motif found in cell cycle control proteins and receptors for cell-cell interactions. *Cell* 60, 981-990.
- Stanisic, T., Sadlowski, R., Lee, C., and Grayhack, J. T. (1978). Partial inhibition of castration induced ventral prostate regression with actinomycin D and cyclohexamide. *Invest. Urol.* 16, 19-22.
- Strasser, A., Harris, A. W., and Cory, S. (1991). *bcl-2* transgene inhibits T cell death and perturbs thymic self-censorship. *Cell* 67, 889-899.
- Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 82, 110-156.
- Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* 100, 64-119.
- Sulston, J. E., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R., and Waterston, R. (1992). The *C. elegans* genome sequencing project: a beginning. *Nature* 356, 37-41.
- Thornberry, N. A., Bull, H. G., Calaycay, J. R., Chapman, K. T., Howard, A. D., Kostura, M. J., Miller, D. K., Molineaux, S. M., Weidner, J. R., Aunins, J., Elliston, K. O., Ayala, J. M., Casano, F. J., Chin, J., Ding, J. F., Egger, L. A., Gaffney, E. P., Limjuco, G., Palyha, O. C., Raju, S. M., Rolando, A. M., Salley, J. P., Yamin, T., Lee, T. D., Shively, J. E., MacCross, M. M., Mumford, R. A., Schmidt, J. A., and Tocci, M. J. (1992). A novel heterodimeric cysteine protease is required for interleukin-1 $\beta$  processing in monocytes. *Nature* 356, 768-774.
- Trent, C., Tsung, N., and Horvitz, H. R. (1983). Egg-laying defective mutants of the nematode *Caenorhabditis elegans*. *Genetics* 104, 619-647.
- Triezenberg, S. J. (1991). Primer extension. In *Current Protocols in Molecular Biology*, F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, eds. (New York: Wiley and Sons), pp. 4.8.1-4.8.5.
- Vaux, D. L., Cory, S., and Adams, J. M. (1988). *Bcl-2* gene promotes haemopoietic cell survival and cooperates with *c-myc* to immortalize pre-B cells. *Nature* 335, 440-442.
- Vaux, D. L., Weissman, I. L., and Kim, S. K. (1992). Prevention of programmed cell death in *Caenorhabditis elegans* by human *bcl-2*. *Science* 258, 1955-1957.
- Way, J. C., and Chalfie, M. (1988). *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. *Cell* 54, 5-16.
- Yochem, J., and Greenwald, I. (1989). *glp-1* and *lin-12*, genes implicated in distinct cell-cell interactions in *C. elegans*, encode similar transmembrane proteins. *Cell* 58, 553-563.
- Yuan, J., and Horvitz, H. R. (1990). Genetic mosaic analyses of *ced-3* and *ced-4*, two genes that control programmed cell death in the nematode *C. elegans*. *Dev. Biol.* 138, 33-41.
- Yuan, J., and Horvitz, H. R. (1992). The *Caenorhabditis elegans* cell death gene *ced-4* encodes a novel protein and is expressed during the period of extensive programmed cell death. *Development* 116, 309-320.
- Yunis, J. J., Frizzera, G., Oken, M. M., McKenna, J., Theologides, and Arnesen, M. (1987). Multiple recurrent genomic defects in follicular lymphoma: a possible model for cancer. *N. Engl. J. Med.* 316, 79-84.