Analysis of the *Caenorhabditis elegans* Axonal Guidance and Outgrowth Gene *unc-33*

Wei Li, Robert K. Herman and Jocelyn E. Shaw

Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota 55108

Manuscript received May 22, 1992
Accepted for publication July 23, 1992

ABSTRACT

Mutations in the *unc-33* gene of the nematode *Caenorhabditis elegans* lead to severely uncoordinated movement, abnormalities in the guidance and outgrowth of the axons of many neurons, and a superabundance of microtubules in neuronal processes. We have cloned *unc-33* by tagging the gene with the transposable element Tc4. Three *unc-33* messages, which are transcribed from a genomic region of at least 10 kb, were identified and characterized. The three messages have common 3′ ends and identical reading frames. The largest (3.8-kb) message consists of the 22-nucleotide *trans*-spliced leader SL1 and 10 exons (I–X); the intermediate-size (3.3-kb) message begins with SL1 spliced to the 5′ end of exon V and includes exons V–X; and the smallest (2.8-kb) message begins within exon VII and also includes exons VIII–X. A γ-ray-induced deletion mutation situated within exon VIII reduces the sizes of all three messages by 0.5 kb. The three putative polypeptides encoded by the three messages overlap in C-terminal sequence but differ by the positions at which their N termini begin; none has significant similarity to any other known protein. A Tc4 insertion in exon VII leads to alterations in splicing that result in three approximately wild-type-size messages: the Tc4 sequence and 28 additional nucleotides are spliced out of the two larger messages; the Tc4 sequence is *trans*-spliced off the smallest message such that SL1 is added 13 nucleotides upstream of the normal 5′ end of the smallest message.

NEARLY 20 years ago, SYDNEY BRENNER (1973, 1974) described the isolation of uncoordinated mutants of *Caenorhabditis elegans* in what he foresaw as the first step in the elucidation of the genetic components of animal behavior. This strategy has indeed proved fruitful. Some of the genes defined by uncoordinated *C. elegans* mutants are now known to encode proteins involved in the assembly and function of the myofilament lattice of body muscle [reviewed by WATERSTON (1988)]. Other uncoordinated *C. elegans* mutants define genes that contribute to the following aspects of nervous system development and function: the cell lineage of neurons or the migration of neuronal precursor cells [reviewed by HORVITZ (1988) and CHALFIE and WHITE (1988)], neurotransmitter function (LEWIS *et al.* 1987; JOHNSON *et al.* 1988; RAND 1989; S. MCINTIRE, E. JORGENSEN and R. HORVITZ, personal communication), synapse formation (HALL and HEDGECOCK 1991), synaptic specificity (J. WHITE *et al.*, unpublished, cited by CHALFIE and WHITE 1988), and axonal outgrowth and guidance [reviewed by HEDGECOCK *et al.* (1987) and CHALFIE and WHITE (1988)] (HEDGECOCK, CULOTTI and HALL 1990; MCINTIRE *et al.* 1992). A gene in the last category is *unc-33*, mutations in which lead to defects in axonal outgrowth and guidance.

As viewed in the dissecting microscope, *unc-33* mutants are severely uncoordinated, almost paralyzed, and move only very slowly. They are also slightly dumpy, partially egg-laying defective (DESAI *et al.* 1988), and defective in defecation (THOMAS 1990).

Defects in neuronal processes in *unc-33* mutants have been seen in several classes of neuron. By Nomarski microscopy, no defects in cell lineage or positions of cell bodies were found for any cells (HEDGECOCK *et al.* 1985). Defects in neuronal processes were first seen by vital staining of certain chemosensory neurons with the dye 5-fluorescein isothiocyanate (FITC). Exposure of wild-type animals to dye results in the staining of six pairs of amphid neurons, in the head, and two pairs of phasmid neurons, in the tail (HEDGECOCK *et al.* 1985). Each stained neuron is bipolar: a ciliated dendritic process extends (anteriorly for the amphids, posteriorly for the phasmids) to a sensillum, and an axonal process extends into the nerve ring, for the amphid neurons, or into the ventral cord, for the phasmid neurons (WHITE *et al.* 1986). Although these neurons stain poorly in *unc-33* mutants, perhaps because of a slight shortening of the sensory cilia (HEDGECOCK *et al.* 1985), it was observed that the axonal processes of both amphids and phasmids tended to terminate prematurely (HEDGECOCK *et al.* 1985). The processes often terminated at abnormal positions and generally had swollen endings. The phasmid axons terminated just as they were to enter the ventral cord. HEDGECOCK *et al.* (1985) also found

defects in the axons of PDE cells, a bilateral pair of neurons that do not fill with FITC in wild-type animals but do fill with dye in either *cat-6* or *che-14* mutants; when one of the latter mutations was used to promote PDE dye filling in an *unc-33* mutant, a variety of axonal abnormalities were observed. In wild-type animals, each PDE neuron extends its axon ventrally into the ventral cord, where it bifurcates and extends both anteriorly and posteriorly in the ventral cord. In *unc-33* mutants, the ventrally directed process was sometimes found to branch prematurely, resulting in a longitudinal axonal branch (or branches) running at a subventral position. Whether running in the ventral cord or not, the anterior branch generally terminated prematurely to give a much shorter process than normal.

DESAI *et al.* (1988) have used immunocytochemical staining with antisera to serotonin to visualize the processes of the pair of hermaphrodite-specific motor neurons (HSNs) by fluorescence microscopy. In wild-type hermaphrodites, a single axon from each of the HSNs, which are necessary for egg-laying, extends ventrally from the HSN cell body into the ventral nerve cord, then turns anteriorly, runs past the vulva, where it usually generates a branch, and extends anteriorly into the nerve ring (WHITE *et al.* 1986; DESAI *et al.* 1988). In *unc-33* hermaphrodites, which are partially egg-laying defective, it was found that the HSN axon generally either made its anteriorly directed turn prematurely, so that it projected anteriorly at a lateral position rather than in the ventral cord, or it stopped prematurely within the ventral cord (DESAI *et al.* 1988). Similar axonal defects of the sensory neurons ALML, ALMR, PLML, PLMR and AVM were reported for *unc-33* animals by SIDDIQUI (1990), who used immunocytochemical staining with an anti-tubulin antibody to see the processes of these cells by fluorescence microscopy. Finally, antibodies to horseradish peroxidase were used to characterize the morphologies of the two pairs of neurons PHC and PVN; four different alleles of *unc-33* were reported to result frequently in misdirected axons of these cells (SIDDIQUI and CULOTTI 1991). All or nearly all neurons whose morphologies have been inspected carefully show axonal defects in *unc-33* mutants. These include sensory neurons, motor neurons and interneurons. It thus seems likely that many other neurons are affected.

HEDGECOCK *et al.* (1985) examined the ultrastructure of the anterior 15 μm of the *unc-33* amphids. No dendritic abnormalities had been revealed in the FITC-filled neurons by fluorescence microscopy (apart from the overall reduction in dye-filling), and the sensilla appeared essentially normal by electron microscopy, too, except for a partial shortening of the cilia of the amphid neurons. The dendrites of apparently all classes of sensory neurons contained a superabundance of microtubules, however. In addition, the microtubules were sometimes larger in diameter than normal or showed abnormal hooks or doublet or triplet tubules. These same defects, superabundance of microtubules and abnormal structures, were observed in the processes of the amphid non-neuronal support cells: the sheath and socket cells; however, the microtubules in muscle and hypodermal cells appeared normal in number and structure. These results led HEDGECOCK *et al.* (1985) to propose that the product of *unc-33* is a component of the axonal cytoskeleton that directs axonal outgrowth and that the axonal guidance defects in the mutants are a consequence of a cytoskeletal defect. In particular, they suggested that the wild-type gene product may be a microtubule-associated protein (MAP) that controls the assembly or stability of neuronal microtubules.

Our aim in the work reported here was to clone and sequence the *unc-33* gene and to characterize the *unc-33* messages in wild type and *unc-33* mutants as a first step in the molecular characterization of the gene and its expression.

MATERIALS AND METHODS

***C. elegans* strains and culture:** Nematodes were grown at 20° on *Escherichia coli* strain OP50 or AMA1004 (CASADABAN *et al.* 1983) on NG agar plates (BRENNER 1974). The mutations *unc-33(e204, e572, e735, e1193* and *mn407)* and *dpy-13(e184)* were derived from the standard Bristol wild-type strain N2. The mutation *unc-33(mn260)* and its derivatives (see below) and *unc-33(rh1030)* and *unc-33(rh1030rh1054)* (provided by E. HEDGECOCK) were derived from strain TR679 (COLLINS, SAARI and ANDERSON 1987).

Isolation of new *unc-33* mutants: The active mutator strain TR679, which exhibits high frequency transposition of several families of transposable elements, was screened visually for spontaneous mutants exhibiting uncoordinated (Unc) phenotypes. One of the Unc mutations, *mn260*, was identified as a new *unc-33* allele by genetic mapping to linkage group *IV* and complementation tests with *unc-33(e204)*. Three independent, spontaneous wild-type revertant strains–bearing *mn260mn325*, *mn260mn327* and *mn260mn329*–were recovered from the original *unc-33(mn260)* mutant strain. The spontaneous reversion frequency was about 10^{-4}. To reduce the number of transposable elements that were TR679-specific but unrelated to the *unc-33* mutation, we crossed the original *unc-33(mn260)* mutant hermaphrodites to N2 males and picked homozygous *mn260* self progeny from the F_1 cross progeny. This procedure was repeated 10 times. After the final outcross, a single *unc-33(mn260)* hermaphrodite was picked to establish the *unc-33(mn260)* strain used for nucleic acid analysis. The *unc-33(mn260mn325)* revertant was outcrossed to +/*unc-33(e204)* males. F_1 cross progeny of genotype *unc-33(mn260mn325)/unc-33(e204)* were recognized by the fact that they segregated Unc self progeny; homozygous *unc-33(mn260mn325)* offspring were picked from these animals. This outcrossing protocol was repeated three times, and the descendants of one *unc-33(mn260mn325)* hermaphrodite were used for subsequent Southern analysis. The other two

revertant strains were outcrossed in the same way.

γ-Ray induced viable *unc-33* mutants were isolated as follows. Wild-type males were exposed to 7,500 roentgens (r) of γ-rays supplied by a ^{137}Cs source at a dose rate of 500 r/min and were mated with *unc-33(e204) dpy-13* hermaphrodites. Unc non-Dpy cross progeny were picked and allowed to self. Putative new *unc-33* mutants were established from F_2 Unc non-Dpy worms that produced only Unc non-Dpy progeny. Viable mutations arose at a frequency of one per 4000 mutagenized haploid genomes. Each new mutant was outcrossed to N2 twice. One of the four new isolates, *mn407*, was shown to have a 0.5-kb deletion in the *unc-33* region, as described in RESULTS; no alterations in *unc-33* DNA structure were detected for the other three mutants.

Nucleic acid analysis: General methods for manipulating DNA and RNA were based on standard procedures (MANIATIS, FRITSCH and SAMBROOK 1982; SAMBROOK, FRITSCH and MANIATIS 1989). Genomic DNA was prepared from animals grown on NG agarose plates by the method of EMMONS, KLASS and HIRSH (1979), with some modifications. Briefly, thoroughly washed worm pellets frozen at −20° were thawed and lysed in 0.1 M Tris-HCl, pH 8.0, 0.05 M EDTA, 0.2 M NaCl, 1% sodium dodecyl sulfate (SDS), and 200 μg/ml proteinase K. The lysate was extracted four times with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), and once with an equal volume of chloroform:isoamyl alcohol (24:1). The DNA was precipitated with two volumes of 95% ethanol and washed twice with 70% ethanol. DNA pellets were dissolved in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0 (TE).

Total RNA was isolated from asynchronous populations of well-fed worms grown on enriched NG agar plates (NG plates with 20 g/liter peptone) seeded with *E. coli* strain AMA1004. Nematodes, washed several times with distilled water, were pelleted and homogenized at 4° in 10 volumes of 4 M guanidine HCl, 0.5% sodium lauryl sarcosinate, 1 mM EDTA, 0.1 M β-mercaptoethanol, and 0.2 M sodium acetate, pH 5.2, with a French pressure cell at 11,000 psi. The homogenates were immediately extracted three times with an equal volume of ice-cold phenol:chloroform:isoamyl alcohol (25:24:1) and once with an equal volume of ice-cold chloroform:isoamyl alcohol (24:1). RNA was precipitated with 0.5 volume of 95% ethanol at −70° overnight. Total RNA was bound to and eluted from an oligo-(dT) cellulose column to enrich for poly(A) RNA (JACOBSON 1987). For size-fractionating RNA, approximately 5 μg of poly(A)-enriched RNA were electrophoresed through formaldehyde gels (1% agarose, 6.6% formaldehyde, 0.04 M 3-[*N*-morpholino]propane sulfonic acid, pH 7.0, 0.05 M sodium acetate, 0.001 M EDTA). Five micrograms of ethidium bromide were added to each of the denatured RNA samples before loading.

Following electrophoresis, DNA was transferred to Zeta-probe Blotting Membrane (Bio-Rad) by capillary blotting with 0.4 M NaOH. Membranes were neutralized in 2 × SSPE (MANIATIS, FRITSCH and SAMBROOK 1982) after transferring. RNA was transferred to Zeta-probe membrane using 20 × SSPE and was fixed to the membranes by baking for 2 hr at 80° under vacuum. ^{32}P-labeled probes were generated by random priming (FEINBERG and VOGELSTEIN 1983). DNA blots were prehybridized in a solution containing 50% formamide, 0.25 M $NaHPO_4$, pH 7.2, 0.25 M NaCl, 7% SDS, 1 mM EDTA, and 100 μg/ml denatured sonicated herring sperm DNA at 42° for more than 6 hr, and hybridized with 10^6 cpm/ml probes in the same solution overnight. The membranes were washed at 50° four times, for 15 min each, in 0.2 × SSPE and 0.1% SDS. For RNA blots, the prehybridization, hybridization and washing solutions were the same as those used for Southern blots, but the temperatures were 50° for prehybridizations and hybridizations and 55° for washing.

Cloning and sequencing the *unc-33* region: Samples of genomic DNA prepared from N2, TR679, an outcrossed *unc-33(mn260)* strain, and three outcrossed revertant strains were digested with several different restriction endonucleases, and the Southern blots were hybridized with probes generated from each of the six *C. elegans* transposons available: Tc1 (EMMONS *et al.* 1983), Tc2 (LEVITT and EMMONS 1989), Tc3 (COLLINS, FORBES and ANDERSON 1989), Tc4 (YUAN *et al.* 1991), Tc5 (provided by J. COLLINS and P. ANDERSON), and Tc6 (DREYFUS and EMMONS 1991). Only the Tc4 probe revealed novel transposon-containing fragments in the *unc-33(mn260)* strain that were absent from all the revertant strains, the parental strain TR679, and N2 (data not shown). When digested with both *Eco*RI and *Pst*I restriction endonucleases and separated on a 0.7% agarose gel, *unc-33(mn260)* genomic DNA exhibited two novel Tc4-containing DNA fragments, 1.7 and 3.8 kb in size. Since Tc4 has two internal *Pst*I sites, these two extra fragments were likely to contain the two ends of a single Tc4 element. To clone the smaller of these two fragments, approximately 100 μg of *unc-33(mn260)* DNA were digested with *Eco*RI and *Pst*I and separated by electrophoresis through an agarose gel. DNA approximately 1.7 kb in size was electrophoresed from the gel onto DEAE-cellulose membrane (Schleicher & Schuell); eluted with 1 M NaCl, 50 mM arginine; ligated with pUC119 plasmid (VIEIRA and MESSING 1987) that had been digested with *Eco*RI and *Pst*I; and transformed into *E. coli* DH5αF′ (HANAHAN 1985). The resulting transformants were transferred to nitrocellulose filters and probed with ^{32}P-labeled Tc4 probe. One recombinant clone, pEP1.7, was found to carry a 1.7-kb DNA fragment that contained a portion of Tc4 and flanking *C. elegans* DNA.

To clone the entire *unc-33* gene, the unique DNA flanking the Tc4 element was separated from the plasmid pEP1.7 by *Eco*RI and *Sac*I double digestion and used as a probe to screen a λEMBL4 N2 genomic library (provided by C. LINK). One hybridizing clone, EH#141, was identified and its 4.2-kb *Eco*RI fragment was shown to hybridize with the probe. The 4.2-kb fragment was subcloned into pUC119 and was used to isolate four more genomic clones (including EH#28) with overlapping inserts from the *unc-33* region. We were unable to detect genomic clones from the λEMBL4 library that extended to the 3′ end of the *unc-33* gene, and no such clones had been described in conjunction with the *C. elegans* genome mapping project (COULSON *et al.* 1986; A. COULSON and J. SULSTON, personal communication). To extend the genomic clones to the 3′ end of the gene, a λ phage library was constructed with DNA from the YAC clone Y37G9, which had been shown to hybridize with EH#28 by R. WATERSTON (personal communication). Stationary phase yeast cells were embedded in 1% solid agarose and lysed as described by SCHWARTZ and CANTOR (1984). The yeast chromosomal DNA was separated through a 1% low melting temperature agarose gel run at 200 V with an 80-sec switch interval for 24 hr at 4° in a CHEF (contour-clamped homogeneous electric field) hexagonal-array apparatus (CHU, VOLLRATH and DAVID 1986). After electrophoresis, the gel was stained with ethidium bromide. A gel slice containing the 460-kb YAC chromosome that had been identified previously by Southern blotting with the 4.2-kb *Eco*RI fragment was isolated, and the agarose was degraded by incubating the slice with agarase (BURMEISTER and LEHRACH 1989). The purified 460-kb YAC chromosome was

then partially digested with *Sau*3A and ligated with λ-DASH (Stratagene) that had been digested with *Bam*HI. Phage were packaged *in vitro* and plated on P2392 (Stratagene). Plaques were probed with a 1.2-kb *Eco*RI fragment from the cDNA clone EH#12-1 (see below). One hybridizing phage, EH#37, was recovered from the library; its insert extended approximately 5 kb beyond EH#28 (see Figure 2) and included the 3′ end of *unc-33*.

Nine cDNA clones were isolated with the 4.2-kb *Eco*RI fragment of EH#141 as probe from two independent λgt10 cDNA libraries: one prepared from early embryonic N2 RNA (provided by J. AHRINGER and J. KIMBLE), the other from mixed-stage N2 RNA (provided by S. KIM and R. HORVITZ). These cDNA clones did not extend to the 3′ end of the transcribed region. A 0.2-kb *Eco*RI cDNA fragment from the 3′ end of one of the cDNA clones, EH#5-1, was used to isolate the tenth clone, EH#12-1, which contained a poly(A) tract. Together, the inserts in these cDNA clones represent a 3.7-kb continuous message from the *unc-33* region.

The sequence of the 3.7-kb cDNA was determined from the three overlapping cDNA clones, EH#1-1, EH#5-1 and EH#12-1. These three clones were chosen as the minimal number needed to cover the region represented by cDNA clones (Figure 2). The entire sequence of EH#1-1 was determined from overlapping subclones and from nested deletions generated by exonuclease III and S1 nuclease digestion (HENIKOFF 1987). The two 0.2-kb *Eco*RI fragments at the 3′ end of EH#5-1 were sequenced separately, and their junction was confirmed by sequencing an overlapping subclone. The 1.2-kb *unc-33* insert from EH#12-1 was sequenced from nested deletions (HENIKOFF 1987). Both cDNA strands were sequenced at least twice. The genomic fragments sequenced were all subcloned from *Eco*RI-digested EH#28 DNA except for the 2.7-kb *Eco*RI fragment at the most 3′ end, which was from EH#37. Genomic sequences were determined from nested deletions, and the sequences of *Eco*RI junctions were confirmed by sequencing subclones that overlapped adjacent *Eco*RI fragments. The 13-kb genomic sequence was determined from only one of the strands. DNA sequencing was performed by the dideoxy chain-termination method (SANGER, NICKLEN and COULSON 1977) with [^{35}S]dATP as the labeled nucleotide, using Sequenase (U. S. Biochemical) as recommended by the manufacturer. Six percent polyacrylamide-8 M urea sequencing gels were run as described by SHEEN and SEED (1988).

Polymerase chain reaction (PCR) amplification of the 5′ ends of the *unc-33* transcripts: The sequences at the 5′ ends of *unc-33* transcripts were determined by performing anchored RACE (rapid amplification of cDNA ends) PCR (FROHMAN, DUSH and MARTIN 1988; FROHMAN 1990). To prepare first-strand cDNA, one μg of poly(A)-enriched RNA prepared from a mixed-stage population of N2 or *unc-33(mn260)* worms was incubated in a 25-μl mixture containing 10 units of avian myeloblastosis virus (AMV) reverse transcriptase in 1 × PCR buffer (Promega Biotech), 6.5 mM $MgCl_2$, 1 mM of each deoxyribonucleoside triphosphate (dNTP), 10 units of RNasin (Promega), and 10 pmol of reverse transcription (RT) primer 5′-RT (5′-CAGCCAGTCGAGAGATCATC-3′) (see Figure 3) at 42° for 1 hr and at 52° for another hour. The cDNA was passed through Centricon-100 spin filters (Amicon Corp.) twice in 2 ml of 0.1 × TE to remove excess 5′-RT primer, and was tailed with poly(A) using dATP and 10 units of terminal deoxynucleotidyl-transferase. The reaction mixture was then diluted with 300 μl of 0.1 × TE. Ten microliters of the diluted cDNA were denatured at 95° for 5 min, annealed at 50° for 3 min in a 50-μl solution containing 1 × PCR buffer, 0.2 mM of each dNTP, 25 pmol of dT17-adaptor primer (5′-GACTCGAGTCGACATCGATTTT-TTTTTTTTTTTTT-3′), 50 pmol of 17-mer adaptor primer (5′-GACTCGAGTCGACATCG-3′), and 50 pmol of the gene-specific amplification (amp) primer 5′-amp (5′-TCGGTGACCTGAGTGTAGAC-3′; see Figure 3). The 5′ ends of the N2 and the *unc-33(mn260)* transcripts were also amplified with 50 pmol of SL1 primer (5′-GGTTTAAT-TACCCAAGTTCGAG-3′) in place of the two adaptor primers. The second cDNA strands were synthesized by adding 2.5 units of Taq polymerase to the above mixtures at 72° and extending at 72° for 30 min. A DNA Thermal Cycler (Perkin-Elmer Cetus) was used to carry out 40 cycles of amplification by running a step-cycle file of 94°, 1 min; 55°, 1 min; 72°, 2 min, followed by a 15 min final extension at 72°. The PCR products were digested with *Cla*I, which cuts once in the adaptor sequence and once within the amplified cDNA immediately 5′ of the 5′-amp primer. Digested products were cloned into *Acc*I-digested pUC119. For PCR products amplified with SL1, the ends were made blunt (HENIKOFF 1987) prior to digestion with *Cla*I, and the fragments were then cloned into *Sma*I- and *Cla*I-digested pUC119. Transformants that hybridized to probes generated from subcloned cDNA fragments upstream of the 5′-amp primer were identified for sequencing.

RNAase protection assay: The position of the 5′ end of the *unc-33* 2.8-kb transcript from N2 and the structures of transcripts from *unc-33(mn260)* animals were analyzed by RNAase protection, essentially as described by MELTON *et al.* (1984). A 0.45-kb N2 cDNA segment (starting at position 1893 and ending at position 3268 as shown in Figure 3) that includes part of exon VI and almost all of exon VII (Figure 6B) was subcloned into the *Hin*dIII site of pGEM-3Zf(+) (Promega) to create pEH.4–5. A ^{32}P-labeled 470-nucleotide antisense RNA probe was generated *in vitro* by transcription, with SP6 RNA polymerase, of the plasmid pEH.4-5 linearized with *Bam*HI. The DNA template was then removed by incubating the transcription mixture with 10 units of RNAase-free DNase I at 37° for 30 min. Approximately 20 μg of total RNA isolated from N2 or *unc-33(mn260)* worms were used for RNA-RNA hybridization at 45° or 60° overnight. Ribonuclease digestion was done at 37° for 45 min. A control reaction using 20 μg of yeast tRNA in place of *C. elegans* RNA was performed under identical conditions. The protected products were resolved on a 6% polyacrylamide-8 M urea sequencing gel.

Sequencing Tc4 junctions in *unc-33(mn260)* and *unc-33(rh1030)*: The right junction (as shown in Figure 2) of the Tc4 insertion in *unc-33(mn260)* was determined by sequencing the recombinant clone pEP1.7, which contains part of the Tc4 element and flanking DNA of *unc-33(mn260)*. Since the left junction was not subcloned, PCR amplification was used to reveal its sequence. Genomic DNA prepared from the outcrossed *mn260* strain was digested with restriction endonuclease *Bgl*II to prevent hairpin formation by Tc4, which has an inverted repeat structure (YUAN *et al.* 1991). Fifty ng of digested DNA were amplified using the *unc-33*-specific primer (5′-AATGGAGACACGCCGACAGA-3′), which is 83 bp from the Tc4 insertion site, and the Tc4-specific primer (5′-CCTGCAGAGATATGGAGCCT-3′), which corresponds to Tc4 nucleotide positions 659–678 defined by YUAN *et al.* (1991). The procedures for amplification and for cloning the amplified genomic fragment were the same as those described above for PCR amplification of the 5′ ends of the *unc-33* transcripts. The 764-bp amplified fragment was blunt-ended, digested with *Cla*I, which cuts within the amplified Tc4 sequence, and then cloned. The desired clone was recognized by hybridization to Tc4, and

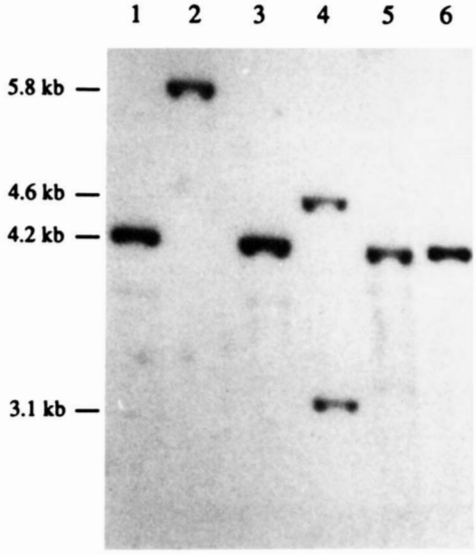


FIGURE 1.—Detection of insertions associated with two spontaneous *unc-33* mutations. Genomic DNA (approximately, 3 μg) was digested with *Eco*RI, electrophoresed through a 0.7% agarose gel and transferred to a nylon membrane. The Southern blot was hybridized with a 4.2-kb *Eco*RI fragment cloned from N2. N2 (lane 1); *unc-33(mn260)* (lane 2)–the hybridizing fragment was larger by 1.6 kb, the size of the Tc4 element; revertant *unc-33(mn260mn325)* (lane 3)–the hybridizing fragment was restored to wild-type size; *unc-33(rh1030)* (lane 4)–two hybridizing fragments of altered size were detected; revertant *unc-33(rh1030rh1054)* (lane 5)–the hybridizing fragment was restored to wild-type size; parental strain TR679 (lane 6).

the ends of its insert were sequenced to obtain the junction of Tc4 in *unc-33(mn260)*.

The entire 3.5-kb Tc4 element inserted in *unc-33(rh1030)*, together with 0.4 kb of flanking *unc-33* DNA, was cloned from the 3.9-kb fraction of *Sac*I-digested *rh1030* genomic DNA, as described for the cloning of pEP1.7 from *unc-33(mn260)* DNA. Recombinant clone pS4-2, which contains the inserted Tc4 element, was identified by probing with the 0.4-kb unique *unc-33* fragment. The junctions of Tc4 in *unc-33(rh1030)* were analyzed by sequencing nested deletions of pS4-2 (HENIKOFF 1987).

Computer analysis: DNA and putative protein sequences were analyzed using IntelliGenetics programs. The FastDB program was used to search the PIR, Swiss-Prot, EMBL and GenBank databanks for amino acid sequences similar to those of the *unc-33* polypeptides.

RESULTS

Identification of a Tc4 element closely associated with *unc-33*: To begin to clone the *unc-33* gene, we sought mutant alleles of *unc-33* that were likely due to insertion of a transposable element within the gene. The *C. elegans* mutator strain TR679 has been shown to exhibit active transposition of Tc1 as well as other families of transposons (COLLINS, SAARI and ANDERSON 1987; COLLINS, FORBES and ANDERSON 1989; YUAN *et al.* 1991); therefore, TR679 was screened visually for spontaneous uncoordinated (Unc) mutants. One Unc mutant that was isolated was shown, by genetic mapping and complementation tests, to carry a new *unc-33* mutation, called *mn260*. In the

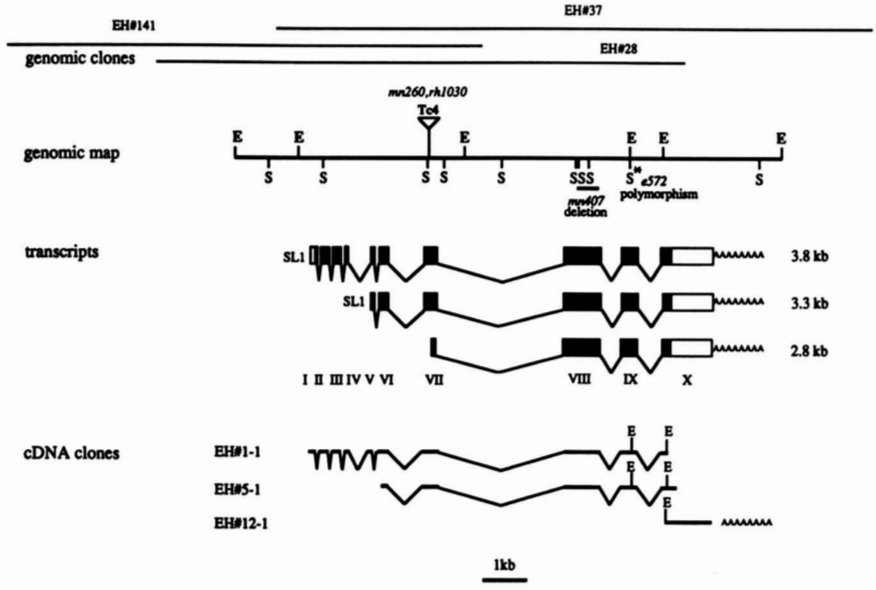
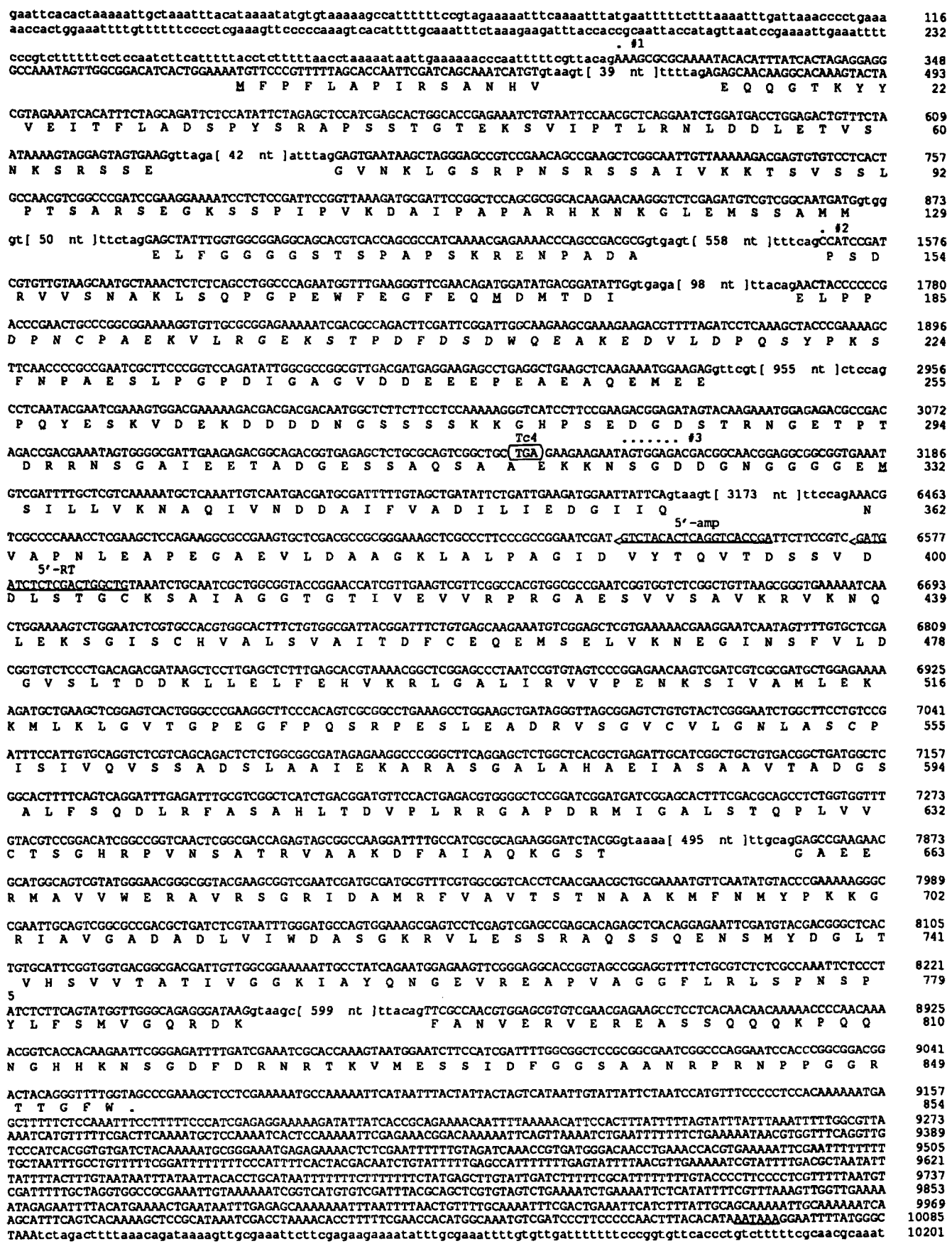


FIGURE 2.—Physical map of the *unc-33* region. A genomic restriction map in the region of *unc-33* is diagrammed for the sequenced 13-kb genomic region covered by three overlapping λ clones shown above the map (E = *Eco*RI and S = *Sac*I). The sites of two Tc4 insertions, *unc-33(mn260)* and *unc-33(rh1030)*, are shown above the genomic map. Below the map are indicated a 0.5-kb deletion, *unc-33(mn407)*, and the *Sac*I site (marked with *) that is absent from the EMS-induced mutation *unc-33(e572)*. The structures of the three *unc-33* transcripts were inferred from the sequences of cDNA clones and PCR clones and from Northern blot analysis. The 3.8- and 3.3-kb transcripts start with a *trans*-spliced leader, SL1; the 2.8-kb transcript has no spliced leader. The filled boxes represent exons containing ORFs, whereas the open boxes represent untranslated regions of exons. The exons are numbered below the transcripts. The three overlapping cDNA clones used to determine the *unc-33* cDNA sequence are shown at the bottom.



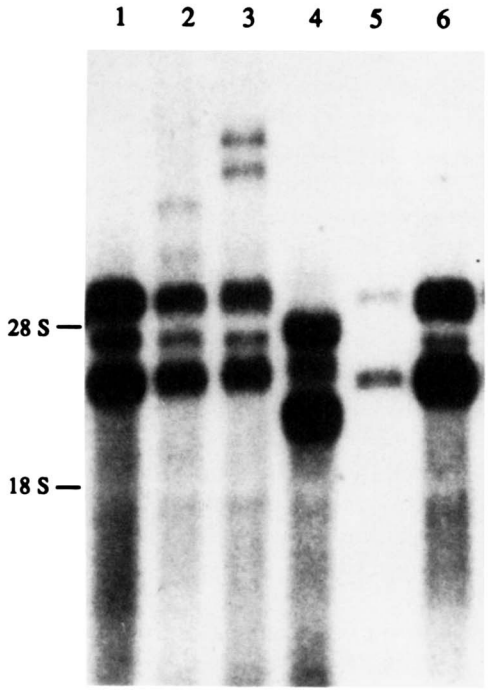


FIGURE 4.—Detection of the *unc-33* transcripts. Poly(A)-enriched RNA was prepared from mixed-stage hermaphrodites: N2, lane 1; *unc-33(mn260)*, lane 2; *unc-33(rh1030)*, lane 3; *unc-33(mn407)*, lane 4; *unc-33(e735)*, lane 5; and *unc-33(e572)*, lane 6. RNA was run through 1% agarose-6.6% formaldehyde gels. The northern blot was hybridized with a probe generated from the 3′ untranslated region of *unc-33*. The two Tc4 mutants exhibited at least two larger transcripts in addition to the three normal-size ones. All three messages from the deletion mutant *unc-33(mn407)* are 0.5 kb smaller than their wild-type counterparts. The two EMS-induced mutants, *unc-33(e735)* and *unc-33(e572)*, have three transcripts indistinguishable from wild type. The bands in the *unc-33(e735)* lane are fainter because less than 1 μg of *unc-33(e735)* RNA was loaded (approximately 5 μg RNA was loaded in each of the other lanes). The film was exposed with an intensifying screen for 10 days at −70°.

mutator genetic background, many insertional mutations tend to revert spontaneously by excision. The *unc-33(mn260)* mutant reverted to wild type at high frequency, and three independent revertant lines were established.

When hybridized with Tc4 probe, Southern blots of genomic DNA extracted from N2, TR679, an outcrossed *unc-33(mn260)* strain, and three outcrossed revertant strains revealed novel Tc4-containing fragments in the *unc-33(mn260)* strain that were absent from all the revertant strains, the parental strain TR679, and N2 (data not shown). In particular, the mutant DNA digested with both *Eco*RI and *Pst*I exhibited two novel Tc4-containing fragments, 3.8 and 1.7 kb in size. Since there are two sites recognized by *Pst*I within Tc4 (YUAN *et al.* 1991), it was very likely that the two new Tc4-containing fragments represented a single Tc4 insertion event associated with the *unc-33* mutation. The smaller Tc4-containing fragment from the *unc-33(mn260)* strain was cloned into pUC119. From this recombinant plasmid, a 0.6-kb *Eco*RI/*Sac*I DNA fragment containing unique worm DNA flanking the Tc4 element was subcloned. The 0.6-kb *Eco*RI/*Sac*I probe detected a 1.7-kb DNA fragment in the mutant and a 4.2-kb fragment in the revertants, TR679, and N2 when used to reprobe the stripped Southern blot of genomic DNA digested with *Eco*RI and *Pst*I (data not shown).

The 0.6-kb unique sequence was subsequently used to screen a *C. elegans* N2 genomic λ library (provided by C. LINK), and one hybridizing clone, EH#141, was identified. A 4.2-kb *Eco*RI fragment from EH#141 that hybridized to the 0.6-kb sequence was used to isolate four more genomic clones, one of which was EH#28. When the 4.2-kb DNA fragment was used to probe a Southern blot with genomic DNA from N2, mutant and revertant strains digested with both *Eco*RI and *Pst*I, a single 4.2-kb hybridizing fragment was observed in N2 and revertant DNA, whereas 3.8- and 1.7-kb fragments were detected in *unc-33(mn260)* DNA. These results confirmed that the 3.8- and 1.7-kb Tc4-containing fragments in *unc-33(mn260)* are derived from a single Tc4 insertion in the 4.2-kb *Eco*RI fragment (data not shown).

Detection of DNA alterations in other *unc-33* mutants: To confirm that the cloned worm sequence did come from the *unc-33* region, we studied another spontaneous *unc-33* mutation, *rh1030*, and its revertant, *rh1030rh1054* (provided by E. HEDGECOCK). DNA extracted from N2, TR679, *unc-33(mn260)*, *unc-33(mn260mn325)*, *unc-33(rh1030)* and *unc-33(rh1030rh1054)* strains was digested with *Eco*RI and probed with the 4.2-kb *Eco*RI fragment (Figure 1). The probe hybridized to a 4.2-kb band in N2, the parental mutator strain TR679, and the revertants of

FIGURE 3.—Genomic nucleotide sequence and predicted amino acid sequences of the *unc-33* gene. The nucleotide sequence shown begins at an *Eco*RI site 377 bp upstream of the putative translational initiator methionine codon of the largest transcript and ends 130 bp downstream of the polyadenylation signal AATAAA (underlined). The numbering of nucleotide positions is indicated to the right of the figure. Amino acid residues are also numbered at the right of the figure, with numbers respresenting residues of the largest predicted polypeptide. Genomic sequences that are not included in cDNA clones are shown in lower case. The sequence data have been submitted to EMBL under accession numbers Z1416 (cDNA sequence) and Z1418 (genomic sequence). The numbers within the [ ] symbols indicate the lengths of the intron sequences that are omitted in the figure. The Tc4 insertion sites for mutations *unc-33(mn260)* and *unc-33(rh1030)*, at position 3139–3141 (the duplicated TGA target flanking the inserted Tc4 elements), is boxed. The 3.8- and 3.3-kb messages start at positions 313 (#1) and 1568 (#2), respectively, and each is preceded by the *trans*-spliced leader SL1. The 2.8-kb message (#3) begins within the range of positions 3152–3159. The transcript start sites are marked with dots above the nucleotides. The putative translational initiator methionine of each of the transcripts is underlined. The primers used in PCR analysis for the initial reverse transcription to produce cDNA (5′-RT) and for the amplification of cDNA (5′-amp) were complementary to the sequences underlined with long arrows.

both mutations. The band was shifted up by 1.6 kb, the size of Tc4 (YUAN *et al.* 1991), in the *unc-33(mn260)* strain. In the *unc-33(rh1030)* strain, however, two bands were observed: 4.6 and 3.1 kb, indicating that the insertion responsible for this mutation is at least 3.5 kb long and has an internal *Eco*RI site. Southern blot analysis of DNA from the *unc-33(rh1030)* strain digested with *Sac*I and hybridized with probes generated from six *C. elegans* transposons indicated that *rh1030* was also caused by a Tc4 element inserted in the *unc-33* region (data not shown). The Tc4 insertion site in *rh1030* was mapped to the same 0.4-kb *Sac*I wild-type fragment in which the Tc4-*mn260* insertion occurred (Figure 2). We sequenced both the left and right junctions of both Tc4-*mn260* and Tc4-*rh1030* and found that the identical TGA trinucleotide in the *unc-33* genomic sequence (Figure 3) abutted each Tc4 terminus. We draw two conclusions: first, that the insertion of each Tc4 involved the duplication of a trinucleotide t, as was previously found for two other Tc4 insertions (YUAN *et al.* 1991), and second, that the two Tc4 insertions in *unc-33* occurred at the identical trinucleotide site. The 3.5-kb Tc4 element in *unc-33(rh1030)* was cloned, and it was found by restriction mapping and Southern blotting that this particular copy of Tc4 has a 1.9–2.0 kb internal fragment that is absent from about two-thirds of the Tc4 copies in *C. elegans*, including Tc4-*mn260* and the two reported by YUAN *et al.* (1991) (W. LI and J. SHAW, unpublished).

Additional evidence supporting the conclusion that cloned sequences were from the *unc-33* region came from the study of other *unc-33* mutations. DNA from N2, the EMS-induced mutations *unc-33(e204, e572, e735, e1193)* and the γ-ray-induced mutation *unc-33(mn407)* was digested with *Sac*I and probed with the genomic clones EH#28 and EH#37 (see below for derivation of EH#37, which was not one of the five original genomic clones). A *Sac*I site was eliminated from the EMS-induced mutation *unc-33(e572)*, and a 0.5-kb sequence was deleted in the γ-ray-induced *unc-33(mn407)* mutant (data not shown). Each of these alterations was located about 4 kb from the Tc4 insertion sites (Figure 2). When this analysis was repeated with *Eco*RI in place of *Sac*I, the 0.5-kb deletion was again recognized for *mn407*, but none of the four EMS-induced mutants was different from N2.

Isolation and sequence analysis of genomic and cDNA clones of *unc-33*: Using as a probe the 4.2-kb DNA fragment of EH#141, nine cDNA clones were isolated from λgt10 libraries provided by J. AHRINGER and J. KIMBLE, and S. KIM and R. HORVITZ. Of these cDNA clones, EH#1-1 and EH#5-1 extended farthest toward the 5′ and 3′ ends, respectively, of the *unc-33* transcript (see Figure 2). A DNA fragment containing the most 3′ region of EH#5-1 was used to identify the cDNA clone EH#12-1, which represented the 3′ end of the *unc-33* transcript, including a poly(A)-tail. Of the genomic clones that were isolated in our original identification of the *unc-33* region, EH#141 and EH#28 represented the most extensive sequences of the gene; however, neither clone extended to the 3′ end of the transcribed region of *unc-33*. The genomic clone EH#141 had failed to identify any cosmids in the collection of A. COULSON and J. SULSTON (personal communication) that might have contained genomic sequences flanking *unc-33* (COULSON *et al.* 1986); therefore, a mini library was constructed from the YAC clone Y37G9 (COULSON *et al.* 1988) that hybridized with EH#28 (R. WATERSTON, personal communication). From this mini genomic library, EH#37 was isolated. EH#37 overlapped EH#141 and EH#28 and extended approximately 4 kb beyond the 3′ end of the *unc-33* transcribed region.

By Southern blotting and sequence analysis, the inserts of three cDNA clones, EH#1-1, EH#5-1 and EH#12-1, were shown to overlap each other. Together their sequences represent a 3.7-kb (excluding the poly(A) tail) cDNA sequence transcribed from the *unc-33* region (Figure 3). The sequence of the first 13 nucleotides at the very 5′ end of the 3.7-kb *unc-33* message (plus a *trans*-spliced leader) was determined by PCR analysis (see below). Thirteen kilobases of *unc-33* genomic sequence were also determined from the inserts of EH#28 and EH#37. The intron/exon structure of the gene was revealed by comparing the cDNA and genomic sequences (Figure 2). The *unc-33* genomic sequence spans at least 10 kb, with 10 exons and 9 introns. The intron sizes ranged from 51 to 3185 bp, and the 5′ and 3′ intron splicing sites were consistent with the worm consensus splicing sequences (EMMONS 1988).

The *unc-33* sequence, including the 3.7-kb cDNA and its deduced amino acid sequence, are shown in Figure 3. The longest putative open reading frame (ORF) would be 2.6 kb, starting from the first ATG near the 5′ end of the cDNA sequence at position 378 and ending at the stop codon TAG at position 9057. There are additional downstream translational stop codons in all three reading frames within 70 nucleotides of this stop. We propose that translational initiation starts at the most 5′ methionine codon, at position 378, because the next in-frame ATG in the cDNA sequence is about 400 bp downstream and the intervening sequence is all open. The 3′ untranslated region is 1.1 kb long with a potential polyadenylation signal AATAAA ending at position 10071, 15–18 nucleotides upstream of the poly(A) tail.

Identification of three *unc-33* transcripts: Poly(A)-enriched RNAs were prepared from mixed-stage populations of N2 and several *unc-33* mutants. The RNAs were analyzed by Northern blotting with a probe from

a portion of the cDNA representing the 3′ untranslated region (Figure 4). In wild-type *C. elegans*, three *unc-33* transcripts, 3.8, 3.3 and 2.8 kb in size, were detected. The relative abundances of the 3.8- and 2.8-kb transcripts were about equal; the abundance of the 3.3-kb transcript was much less than the other two. All three transcripts were 0.5 kb smaller in the *unc-33(mn407)* deletion mutant, strongly indicating that all three RNA species are transcribed from the *unc-33* region. Of the two EMS mutants tested, *unc-33(e572)* and *unc-33(e735)*, neither differed from N2 by Northern blot analysis. [The *unc-33(e735)* lane was much fainter than the other lanes because less RNA was loaded.] In the two Tc4 insertional mutants, there were at least two larger transcripts, corresponding to Tc4 insertions in exon VII; however, to our surprise, there were also three normal-sized transcripts. The abundances of the Tc4-containing transcripts were much lower than those of the normal-sized ones, which seemed to be about as abundant as the corresponding N2 transcripts.

By Northern blot analysis with probes made from individual exons of *unc-33*, the three wild-type transcripts were found to differ only at their 5′ ends (data not shown; see Figure 2). The probes generated from the first four exons, I-IV, appeared to hybridize to the 3.8-kb transcript only. Probes from exons V and VI hybridized to both of the larger transcripts but not to the 2.8-kb transcript. When probe from exon VII was used, the 2.8-kb band on the Northern blot was apparent but of very low intensity, suggesting that the 2.8-kb transcript has its 5′ end within exon VII. Probes from exons VIII–X hybridized to all three transcripts. Hybridization with single-stranded RNA probes on Northern blots indicated that all three transcripts are transcribed from the same strand of *unc-33* (data not shown).

Analysis of 5′ ends of wild-type *unc-33* transcripts: Anchored PCR amplification was employed to characterize more precisely the 5′ ends of the *unc-33(+)* transcripts. The primer designed for PCR amplification was located in exon VIII, so that the amplified products would cover the 5′ ends of all three messages. The expected lengths of the PCR products representing the three messages were 1.2, 0.7 and 0.2 kb (corresponding to the 3.8-, 3.3- and 2.8-kb messages, respectively). PCR products of these sizes were identified, cloned and sequenced. Both the 1.2- and 0.7-kb PCR products had the 22-nucleotide sequence of the *trans*-spliced leader SL1 at their 5′ ends (KRAUSE and HIRSH 1987). The sequence of the 1.2-kb product immediately following the SL1 terminus matched the genomic sequence beginning 13 bp 5′ of the sequence obtained from the cDNA clone EH#1-1, *i.e.*, we extended the sequence of exon I 13 bp in the 5′ direction as a consequence of the PCR analysis. The genomic sequence at this position shows a good 3′ splice acceptor sequence (TTACAG|A) for splicing of the SL1 to the 3.8-kb message (Figure 3). The sequence of the 0.7-kb product immediately following its SL1 terminus matched the beginning of exon V, *i.e.*, the same 3′ splice acceptor site used for *cis*-splicing of the 3.8 kb message (at the exon-IV-exon V junction) is used for the *trans*-splicing of SL1 to the 5′ end of the 3.3-kb message. The 0.2-kb PCR product representing the 2.8-kb mRNA (smallest message) did not have a *trans*-spliced leader sequence. Furthermore, PCR amplification reactions using the identical 3′ primer but SL1 as the 5′ primer failed to amplify any 0.2-kb product. Many 0.2-kb anchored PCR clones were sequenced, and most of them started at position 3152 in exon VII (Figures 3 and 5).

RNAase protection assays were subsequently performed to test whether position 3152 represents the real initiation site for the 2.8-kb message or might be the result of reverse-transcription or amplification artifacts (owing, for example, to RNA secondary structure in the region). An RNA probe complementary to mRNA sequences from the middle of exon VI to near the end of exon VII was used in these assays. The observed and expected results of the RNAase protection experiments are shown in Figure 6. From the PCR results, N2 total RNA was expected to protect two fragments: a 409-nucleotide fragment protected by the 3.8- and 3.3-kb transcripts and a 117-nucleotide fragment protected by the 2.8-kb transcript. Two prominent fragments were found in the autoradiogram: one at the same position as a 423-nucleotide single-stranded DNA marker and the other at the position of a 122-nucleotide single-stranded DNA marker. Since RNA samples run approximately 90–95% as fast as single-stranded DNA markers of the same length (SAMBROOK, FRITSCH and MANIATIS 1989), the sizes of the two protected fragments would be around 381–402 and 110–116 nucleotides, respectively, which agree very closely with the expected results.

Analysis of *unc-33(mn260)* transcripts: The wild-type-size transcripts of *unc-33(mn260)* were also analyzed by anchored PCR and RNAase protection experiments (Figures 5 and 6). PCR analysis of the 3.8- and 3.3-kb transcripts revealed that the Tc4 insertion was entirely spliced out, together with 28 nucleotides of exon VII sequence upstream of the Tc4 sequence (three of these nucleotides are those duplicated upon Tc4 insertion). Only one clone was sequenced, so we have not excluded the possibility that other spliced products were also formed. Inspection of the sequence at the 5′ splice site indicates that ACG|GUGAGAG, located 28 nucleotides upstream of the Tc4 element, was a cryptic 5′ splice donor and that the downstream terminus of the Tc4, CCCTAG, provided the 3′ splice

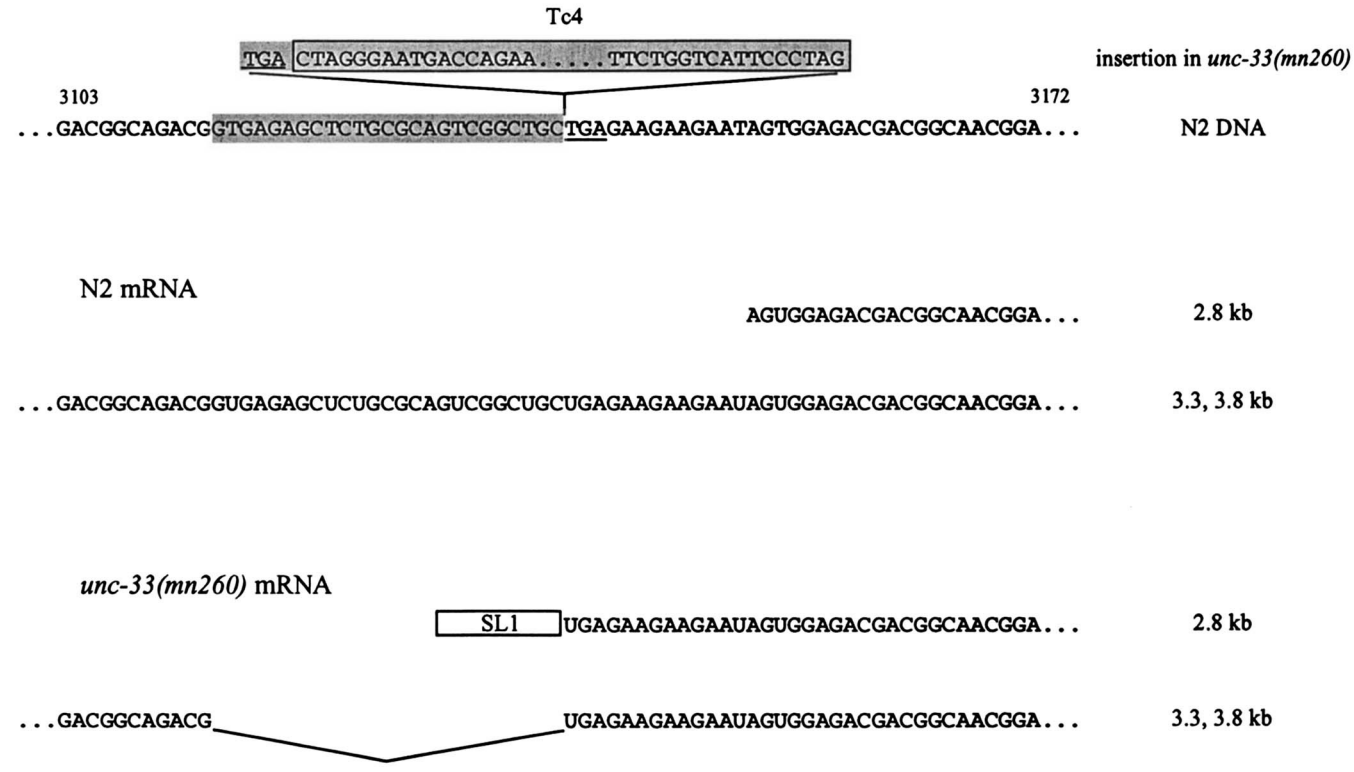


FIGURE 5.—Structure of *unc-33* mRNAs in a portion of exon VII as determined by PCR analysis. The genomic sequence from nucleotide 3103 to nucleotide 3172 of N2 DNA (wholly within exon VII; see Figures 2 and 3) is diagrammed at the top of the figure. The Tc4 element inserted in *unc-33(mn260)* is shown above the N2 sequence. The inserted Tc4 is boxed, and the three base pairs duplicated upon Tc4 insertion are underlined. The sequence corresponding to an intron removed in *unc-33(mn260)* mRNA is shaded. For N2 mRNA, cDNA was synthesized using the 5′-RT primer (Figure 3) and tailed with poly(A). cDNA was amplified using Td17- and 17-mer-adaptor primers (see MATERIALS AND METHODS) and the 5′-amp primer (Figure 3). The sequences through this region of cloned amplification products that correspond to the smallest (2.8 kb) and two largest (3.8, 3.3 kb) mRNAs are illustrated. The 5′ end shown for the N2 2.8-kb transcript was the most frequently observed amplification product; however, amplification products that were a few base pairs longer or shorter were also detected. For *unc-33(mn260)* mRNA, cDNA was synthesized using the 5′-RT primer. cDNA was amplified using the SL1 primer (see MATERIALS and METHODS) and the 5′-amp primer (Figure 3). The sequence through this region of cloned amplification products that correspond to the smallest (2.8 kb) and the two largest (3.3, 3.8 kb) mRNAs are illustrated. The 2.8-kb mRNA of *unc-33(mn260)* is *trans*-spliced to SL1 at the site of Tc4 insertion. The position of the intron removed from the 3.3- and 3.8-kb mRNAs of *unc-33(mn260)* is illustrated by a gap in the sequence.

acceptor site (CCCUAG|U). PCR amplification with the SL1 primer as the 5′ end primer showed that the 5′ end of the 2.8-kb transcript had an SL1 *trans*-spliced leader sequence (which is not present at the 5′ end of the wild-type 2.8-kb transcript). The SL1 sequence of the *unc-33(mn260)* 2.8-kb transcript was followed by sequence that is immediately 3′ of the Tc4 insertion site. Thus, the 2.8-kb transcript, which in wild type initiates about 13 bp downstream of the Tc4 insertion site, most likely initiated within the Tc4 element and obtained the SL1 leader when the Tc4 sequence was spliced out of the precursor RNA; thus, the *trans*-splicing of SL1 used the same 3′ splice acceptor, at the end of Tc4, as was used in the *cis*-splicing of the *unc-33(mn260)* 3.8- and 3.3-kb transcripts.

RNAase protection assays of *unc-33(mn260)* RNA were consistent with these PCR results (Figure 6). Because of the 25-nucleotide deletion in exon VII, protection by the 3.8- and 3.3-kb RNAs would be expected to produce two fragments, 254 and 130 nucleotides long, from the same RNA probe that was used to analyze wild-type *unc-33* messages. Protection by the 2.8-kb RNA would be expected to give rise to a 130-nucleotide fragment. The results agreed well with these expectations; two prominent protected fragments were seen, corresponding to fragments 239–252 and 125–132 nucleotides long. A fragment corresponding to a 423-nucleotide DNA marker was also observed, probably representing incomplete RNAase digestion of RNA hybridizing to the 3.8- and 3.3-kb transcripts. The larger Tc4-containing transcripts present in *unc-33(mn260)* could also potentially protect the probe used in these RNAase protection experiments, producing protected fragments 277–280 and 130–127 nucleotides long. Since the larger Tc4-containing fragments appear to be in very low abundance compared to the wild-type-sized transcripts, they would be expected to produce faint protected fragments. A faint fragment approximately

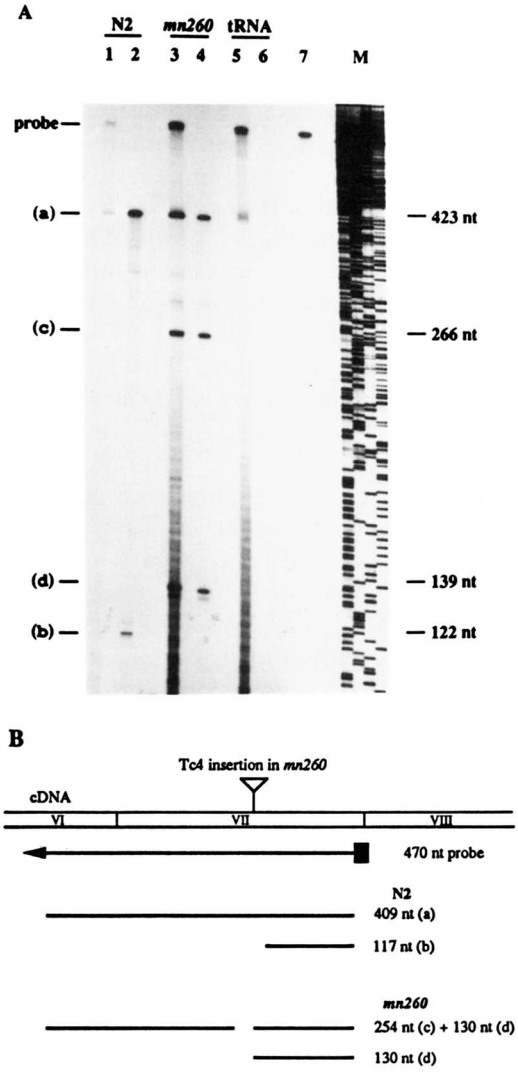


FIGURE 6.—RNAase protection analysis of *unc-33* RNA. (A) Results of an RNAase protection assay of RNA from N2 (lanes 1 and 2), *unc-33(mn260)* (lanes 3 and 4) and yeast tRNA as negative control (lanes 5 and 6); an undigested probe (lane 7). Lanes 1–6 had 20 μg of total RNA. The RNA-RNA hybridizations were performed at 45° (lanes 1, 3 and 5) or 60° (lanes 2, 4 and 6) overnight (the hybridizations conducted at the higher temperature led to clearer results). The protected products were resolved on a sequencing gel. Single-stranded DNA ladders from a sequencing reaction were used as molecular size markers (M). The numbers on the right side of the figure are the sizes of marker DNA at the positions of the protected RNA fragments; RNA molecules run 90–95% as fast as the single-stranded DNA markers of the same length. The letters on the left correspond to the letters given in B. (B) A segment of the *unc-33* cDNA including exon VII and parts of exons VI and VIII is represented diagrammatically at the top. The site of Tc4 insertion in *unc-33(mn260)* is indicated above the diagram. A 470-nucleotide (nt) antisense RNA probe was generated by *in vitro* transcription with SP6 RNA polymerase of pEH.4–5 linearized with *Bam*HI. Sixty-three nucleotides at the 5′ and 3′ ends of this probe are vector sequences not expected to hybridize with *unc-33* RNA. In N2, the two larger transcripts (3.8 and 3.3 kb) are expected to protect a 409-nt fragment (a), and the smallest transcript (2.8 kb) is expected to protect a 117-nt fragment (b). In *unc-33(mn260)*, protection of the probe by the 3.8- and 3.3-kb transcripts, which are missing 25 nucleotides from exon VII, is expected to generate two fragments, 254 nt (c) and 130 nt (d). The 2.8-kb RNA from *unc-33(mn260)*, which begins (after a *trans*-spliced SL1) at the Tc4 insertion site, is expected to protect a 130-nt fragment (d).

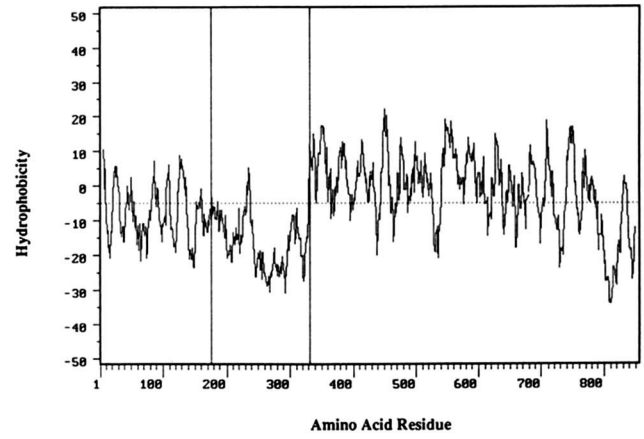


FIGURE 7.—Hydrophobicity profile of the hypothetical *unc-33* polypeptides. The hydropathy index was computed using the SOAP program (PC/Gene program package from IntelliGenetics, Inc.) based on parameters of KYTE and DOOLITTLE (1982), with an 11-residue window. The vertical lines are drawn at residues 176 and 332 to indicate the predicted initiator methionine residues for the putative polypeptides translated from the 3.3- and 2.8-kb messages, respectively.

266–281 nucleotides long (visible just above the 254-nucleotide fragment) could correspond to the 277–280-nucleotide fragment expected to be protected by the Tc4-containing messages.

The *unc-33* gene encodes novel proteins: The longest ORF of the 3.8-kb message is 2.6 kb long and could encode a polypeptide of 854 amino acid residues (Figure 3), with a mass of 90.8 kD. The predicted *unc-33* polypeptide does not reveal any significant amino acid sequence similarity to any other known protein in the databanks, nor does it contain the microtubule-binding domains found in MAP2, tau proteins, or MAP1B (LEWIS, WANG and COWAN 1988; NOBLE, LEWIS and COWAN 1989). Motif analysis indicated that the protein could undergo several kinds of post-translational modification, such as phosphorylation. Calculation of the hydropathy index (KYTE and DOOLITTLE 1982) indicates that the polypeptide has no significant signal peptide and probably no transmembrane domain, suggesting it is a cytoplasmic protein (Figure 7). The 331 amino acid residues at the N terminus of the protein are very hydrophilic, and the region from approximately amino acid residue 180–331 is very negatively charged; the middle segment is more hydrophobic with a weakly predicted transmembrane domain from residue 545–561; the 67 residues at the C terminus are more hydrophilic. The overall pI of the predicted protein is 4.76.

The 3.3-kb message would contain a 2.1-kb ORF if either the first or second in-frame AUG codon of the transcript, located at positions 1640 and 1646, respectively (Figure 3), is the translational initiation site. We tentatively propose that the AUG at position 1640 is the initiator site, since the majority of eukaryotic messages are translated from their first AUG codon

(KOZAK 1986); however, the AUG codon at position 1646 may have a more favorable context for translational initiation, with a purine at position −3 (KOZAK 1986). The predicted polypeptide would be 679 residues long with a mass of 72.1 kD. The N terminus of this polypeptide begins with the highly negatively charged region (described above as residues 180–331 of the largest polypeptide); the calculated pI of the polypeptide is 4.51.

The translational initiator for the 2.8-kb message is most likely located at the first AUG of the transcript at position 3185 (Figure 3). The second in-frame AUG is about 400 nt downstream, and the intervening sequence is all open. The third polypeptide would consist of 523 amino acid residues with a mass of 55.5 kD. As shown in Figure 7, this putative protein differs from the other two in its charge distribution: the entire highly negatively charged N-terminal region is missing, and the pI, 6.29, is consequently much higher. All three predicted polypeptides are translated from the same ORF. No other long ORFs were found.

DISCUSSION

We have identified the *unc-33* region by correlating a Tc4 insertion with an *unc-33* mutation: three independent phenotypic reversions of the mutation were associated with the loss of the inserted Tc4 element. An independent *unc-33* mutation was correlated with the presence of a second Tc4 element, distinguishable from the first but inserted at the same site as the first; reversion of this mutation was also correlated with loss of the inserted element. A third *unc-33* mutation, identified following gamma-ray treatment, was found to be associated with a 0.5-kb deletion in the region, and an EMS-induced *unc-33* allele was associated with an altered *Sac*I restriction site. All four of these mutations lie within what we conclude from our cDNA analysis are exons of the *unc-33* gene.

Three *unc-33* messages were identified and characterized. The three messages have common 3′ ends and identical reading frames; they differ only at their 5′ ends. The largest (3.8 kb) message comprises the 22-nucleotide *trans*-spliced leader SL1 (KRAUSE and HIRSH 1987) and 10 exons (I–X); the intermediate-size (3.3 kb) message begins with SL1 spliced to the 5′ end of exon V and includes exons V–X; and the smallest (2.8 kb) message begins within exon VII and also includes exons VIII–X. The 0.5-kb deletion mutation, *unc-33(mn407)*, is situated wholly within exon VIII, which is common to all messages, and as expected, all three messages in the mutant are reduced in size by 0.5 kb, as judged by northern analysis. It seems likely from the molecular evidence that *mn407* is a null allele, and consistent with this view is the fact that the phenotype conferred by *mn407* is the most severe uncoordination found among different *unc-33* mutants. The *mn407* mutants are viable and fertile, however, which leads us to suggest that *unc-33* is not required for viability or fertility. Both of the Tc4 insertion mutations confer a slightly weaker phenotype than does the *unc-33(mn407)* mutation. The Tc4 insertions occurred at a position slightly 5′ of the beginning of the smallest message. We have analyzed in detail the transcripts produced by *unc-33(mn260::Tc4)*. In addition to a small amount of large transcripts that presumably arose as a consequence of transcription through the inserted Tc4, we found three wild-type size transcripts. The larger two of these, presumably identical to the two larger wild-type mRNAs in their initiation, had the Tc4 insertion plus 28 nucleotides (including three nucleotides duplicated with the insertion of Tc4) spliced out. Both of the resulting transcripts thus contained a 25-nucleotide deletion of exon VII, and hence a frameshift, in their coding sequences compared to the wild-type transcripts. The smallest *unc-33(mn260)* transcript differed from wild type in having an SL1 *trans*-spliced leader plus about 13 nucleotides added to the 5′ untranslated region. Since there are no ATG translational initiation codons within the additional nucleotides, we presume that translation of this transcript would be essentially normal. Translation of this message may account for the slightly hypomorphic phenotype conferred by *unc-33(mn260)*.

Each of the two larger *unc-33(+)* transcripts may have its own promoter, each generating a transcript that is *trans*-spliced to SL1. Alternatively, it is possible that a single promoter is used for both transcripts, in which case the smaller transcript would be formed as an alternatively spliced product, *i.e.*, the 3′ splice site immediately preceding exon V would be used either for *cis*-splicing to exon IV (contributing to the 3.8-kb message) or would be *trans*-spliced to SL1 to generate the 3.3-kb message. The intron between exons IV and V is large (570 nucleotides), and the ratio of *trans*-spliced to *cis*-spliced message (at this 3′ splice site) is low. Sites of transcriptional initiation were difficult to pinpoint because of the *trans*-splicing of the messages; no obvious TATA boxes were found immediately upstream of either of the two larger messages. A consensus promoter with CAT box and TATA box was found 1.1 kb upstream of the beginning of the 3.8-kb message; however, the significance of these sequences is not clear. We believe that the smallest (2.8 kb) message must have its own promoter, because it is initiated within an exon of the other messages. No TATA box was found upstream of the transcriptional initiation site, however. No repeated sequences that would be candidates for transcriptional regulatory sites were found by comparisons of sequences upstream of the transcripts. The only unusual repeated sequence that was noted is located in the 3′

untranslated region of the transcripts. A 42-nucleotide sequence starting at position 9374 has its inverse complement repeated, with only a single mismatch, starting at position 9445. These sequences could form a stem-loop structure in the mRNA with a 42-bp stem and a 28-nucleotide loop.

The splicing patterns found in the region of the Tc4 insertion for the *unc-33(mn260)* mutant were intriguing. Because the Tc4 insertion was only about 13 bp upstream of the wild-type initiation site for the 2.8-kb message, we surmise that the promoter for the 2.8-kb message was separated from its normal transcriptional initiation site by the inserted Tc4 and that transcription was initiated within the Tc4 element. The splicing pattern found for the 3.3- and 3.8-kb *unc-33(mn260)* messages indicates that a 3′ splice site is situated at the 3′ end of the Tc4 element and that no 5′ splice site is present within the Tc4 sequence (*i.e.*, the region spanned by a cryptic 5′ splice site within the *unc-33* gene, upstream of the Tc4 insertion, and a 3′ splice site at the very 3′ end of the Tc4 sequence was spliced out). The work of CONRAD *et al.* (1991) has led to the view that in *C. elegans* a message with no 5′ splice site between the 5′ end of the message and the first 3′ splice site, a segment referred to as an outron, should be *trans*-spliced. It may also be important for the outron to be A-U rich (CONRAD *et al.*, 1991); Tc4 has 67% thymine and adenine (YUAN *et al.* 1991). We therefore would expect, as was found, that the transcript initiated within Tc4 would be *trans*-spliced to SL1 using the same 3′ splice site as was used for the *cis*-splicing of the larger transcripts.

It is curious that one end of the Tc4 element serves as a 3′ splice site. Although the presence of some larger transcripts in *mn260* animals indicates that Tc4 is not always spliced out or is spliced out somewhat slowly, most of the messages that are present have in fact undergone splicing at the 3′ splice site at the end of Tc4, even though, apart from the consensus AG at the intron/exon boundary, there is a poor match to the consensus 3′ splice site (EMMONS 1988). Because Tc4 has long terminal inverted repeats (YUAN *et al.* 1991), a 3′ splice site should be encountered at the end of Tc4 regardless of the direction of transcription. It has been shown that different *Ds* elements inserted within exons of the waxy locus of maize can be spliced out during transcription of the waxy gene, and that the splice sites that are used can be either within and near the ends of the transposon or within the waxy gene itself (WESSLER 1991). This has led to the hypothesis that transposons able to behave as introns would have a strong selective advantage because they would be able to insert into genes with little deleterious effect on gene expression. Tc4 is a rather poor intron, however, as it must rely on a cryptic 5′ splice site, and only certain favorably placed 5′ splice sites would be expected to be harmless for gene expression (of course we would not have identified an *unc-33* mutation in such cases).

The Tc4 transposon family was identified and characterized by YUAN *et al.* (1991), who described two Tc4 insertions, in *unc-86* and *ced-4*. Both mutations occurred spontaneously in the mutator strain TR679, as did the two Tc4 insertions we have described in *unc-33*. The *ced-4* mutation was unstable in the TR679 genetic background (YUAN *et al.* 1991), as were both of the Tc4 insertions in *unc-33* that we have studied. It seems clear that TR679 promotes the transposition of Tc4 (YUAN *et al.* 1991), as well as other apparently unrelated transposons (COLLINS, SAARI and ANDERSON 1987; COLLINS, FORBES and ANDERSON 1989). The two insertions (of two different Tc4 elements) at the identical site within *unc-33* suggests that the site is a hot spot for Tc4 insertion. The trinucleotide target for the insertions in *unc-33* was identical to that for the insertion in *unc-86*, TGA. Indeed, if we include the five bases on each side of the TGA target, then the *unc-33* target matches the *unc-86* target at 11/13 bases. The same comparison between *unc-33* and *ced-4* gives only 7/13 matches (in either orientation). The *unc-33* target matches exactly the consensus CTNAG target suggested by YUAN *et al.* (1991) from their inspection of the *unc-86* and *ced-4* targets.

According to our analysis, each *unc-33* message can encode a distinct polypeptide (523-, 679- and 854-amino acid residues); each of the two smaller polypeptides is expected to be identical in sequence to a C-terminal segment of the largest polypeptide. Indeed, when polyclonal antibodies raised against *unc-33* fusion proteins made from a 2.1-kb cDNA fragment covering all three messages were used in Western blots, three proteins were detected (W. LI and J. SHAW, unpublished). By computer analysis, the three predicted polypeptides did not show significant amino acid sequence similarity to any other peptides in the databanks. Calculation of the hydrophobicity index indicated that the N-terminal third and the C terminus of the largest polypeptide are very hydrophilic and that the middle segment is much more hydrophobic. Interestingly, the highly negatively charged N-terminal region of the largest polypeptide is the extreme N terminus of the middle-sized polypeptide and is missing from the smallest polypeptide.

As noted in the Introduction, different mutant alleles of *unc-33* have been shown to be defective in the guidance and outgrowth of the axons of all or nearly all neurons that have been inspected. The mutations are all recessive to the wild-type allele and by genetic and now molecular criteria are likely to be loss-of-function mutations. On the basis of the observation that neuronal microtubules were superabundant and structurally abnormal in the anterior sensory den-

drites of an *unc-33* mutant that was analyzed by electron microscopy, HEDGECOCK *et al.* (1985) suggested that the *unc-33* gene encodes a component of the neuronal cytoskeleton. In particular, it was suggested that the *unc-33* product may be a microtubule associated protein (MAP). Computer analysis of the putative *unc-33* proteins did not detect any domain homologous to the microtubule-binding domains identified in MAP2, tau proteins or MAP1B. More and more MAP genes have been identified and sequenced, and not all of them have common domains for microtubule binding (WICHE, OBERKANINS and HIMMLER 1991); therefore, it is still possible that the *unc-33* proteins have an uncharacterized microtubule-binding region. Another possibility is that the *unc-33* products affect some other cytoskeletal component, which in turn affects neuronal microtubule number and structure. Interactions between microtubules and intermediate filaments may be mediated by proteins that are known to bind to both of these components (MIYATA *et al.* 1986); other proteins bind microtubules and actin filaments (GRIFFITH and POLLARD 1978). In any case, the hypothesis that *unc-33* affects the neuronal cytoskeleton suggests that the axonal guidance defects in the *unc-33* mutants are a secondary consequence of defects in axonal outgrowth rather than a consequence of misreading extracellular guidance cues. According to this view, *unc-33* proteins would be expected to be intracellular; the *unc-33* coding sequences support this view. We would also expect that the mistakes in axonal guidance or branching and the premature terminations in axonal outgrowth found in *unc-33* animals would all be cell autonomous. This prediction could be tested by analysis of *unc-33* genetic mosaics (HERMAN 1989). Indirect immunofluorescence staining with antibodies to *unc-33* protein has revealed that the *unc-33* proteins are distributed exclusively within neuronal processes after early embryogenesis (W. LI and J. SHAW, unpublished) and supports the notion the *unc-33* function is required in neurons. Further investigations will be required to determine whether the *unc-33* proteins are in fact associated with the neuronal cytoskeleton, to assess the functional role *unc-33* plays in developing *C. elegans* neuronal processes, and to discover whether any *unc-33* homologs are involved in axonal outgrowth and guidance in other animals.

We thank E. HEDGECOCK for *unc-33(rh1030)* and *unc-33(rh1030rh1054)*, R. WATERSTON for identifying and providing YACs that overlap *unc-33*, P. ANDERSON, J. COLLINS, D. DREYFUS, S. EMMONS, R. HORVITZ, A. LEVITT and J. YUAN for transposon probes, C. LINK, J. AHRINGER, J. KIMBLE, S. KIM and R. HORVITZ for genomic and cDNA libraries, J. HODGKIN for strains, A. COULSON and J. SULSTON for help in locating *unc-33* on the physical map, and B. MAGEE and W. CHU for help with the CHEF gels. This work was supported by U.S. Public Health Service grant GM22387 (awarded to R.K.H.) and National Science Foundation grant DCB-8711133 and U.S. Public Health Service grant HD22164 (awarded to J.E.S.).

LITERATURE CITED

BRENNER, S., 1973 The genetics of behaviour. Br. Med. Bull. **29:** 269–271.

BRENNER, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics **77:** 71–94.

BURMEISTER, M., and H. LEHRACH, 1989 Isolation of large DNA fragments from agarose gels using agarase. Trends Genet. **5:** 41.

CASADABAN, M. J., A. MARTINEZ-ARIAS, S. K. SHAPIRA and J. CHOU, 1983 β-Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Methods Enzymol. **100:** 293–308.

CHALFIE, M., and J. WHITE, 1988 The nervous system, pp. 337–391 in *The Nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

CHU, G., D. VOLLRATH and R. W. DAVIS, 1986 Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science **234:** 1582–1585.

COLLINS, J., E. FORBES and P. ANDERSON, 1989 The Tc3 family of transposable genetic elements in *Caenorhabditis elegans*. Genetics **121:** 47–55.

COLLINS, J., B. SAARI and P. ANDERSON, 1987 Activation of a transposable element in the germline but not the soma of *Caenorhabditis elegans*. Nature **328:** 726–728.

CONRAD, R., J. THOMAS, J. SPIETH and T. BLUMETHAL, 1991 Insertion of part of an intron into the 5′ untranslated region of a *Caenorhabditis elegans* gene converts it into a *trans*-spliced gene. Mol. Cell. Biol. **11:** 1921–1926.

COULSON, A., J. SULSTON, S. BRENNER and J. KARN, 1986 Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **83:** 7821–7825.

COULSON, A., R. WATERSTON, J. KIFF, J. SULSTON and Y. KOHARA, 1988 Genome linking with yeast artificial chromosomes. Nature **335:** 184–186.

DESAI, C., G. GARRIGA, S. L. MCINTYRE and H. R. HORVITZ, 1988 A genetic pathway for the development of *Caenorhabditis elegans* HSN motor neurons. Nature **336:** 638–646.

DREYFUS, D. H., and S. W. EMMONS, 1991 A transposon-related palindromic repetitive sequence from *C. elegans*. Nucleic Acids Res. **19:** 1871–1877.

EMMONS, S. W., 1988 The genome, pp. 47–79 in *The Nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

EMMONS, S. W., M. R. KLASS and D. HIRSH, 1979 Analysis of the constancy of DNA sequences during development and evolution of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **76:** 1333–1337.

EMMONS, S. W., L. YESNER, K.-S. RUAN and D. KATZENBERG, 1983 Evidence for a transposon in Caenorhabditis elegans. Cell **32:** 55–65.

FEINBERG, A. P., and B. VOGELSTEIN, 1983 A technique for radiolabelling DNA restriction fragments to high specific activity. Anal. Biochem. **132:** 6–13.

FROHMAN, M. A., 1990 RACE: rapid amplification of cDNA ends, pp. 28–37 in *PCR Protocols: A Guide to Methods and Applications*, edited by M. A. INNIS, D. H. GELFAND, J. J. SNINSKY and T. J. WHITE. Academic Press, New York.

FROHMAN, M. A., M. K. DUSH and G. R. MARTIN, 1988 Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85:** 8998–9002.

GRIFFITH, L. M., and T. D. POLLARD, 1978 Evidence for actin

filament-microtubule interaction mediated by microtubule-associated proteins. J. Cell Biol. **78:** 958–965.

HALL, D. H., and E. M. HEDGECOCK, 1991 Kinesin-related gene *unc-104* is required for axonal transport of synaptic vesicles in *C. elegans*. Cell **65:** 837–847.

HANAHAN, D., 1985 Techniques for transformation of *Escherichia coli*, pp. 109–135 in *DNA Cloning*, Vol 1, edited by D. M. GLOVER. IRC Press, Washington, D.C.

HEDGECOCK, E. M., J. G. CULOTTI and D. H. HALL, 1990 The *unc-5*, *unc-6*, and *unc-40* genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. Neuron **4:** 61–85.

HEDGECOCK, E. M., J. G. CULOTTI, J. N. THOMSON and L. A. PERKINS, 1985 Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes. Dev. Biol. **111:** 158–170.

HEDGECOCK, E. M., J. G. CULOTTI, D. H. HALL and B. D. STERN, 1987 Genetics of cell and axon migration in *Caenorhabditis elegans*. Development **100:** 365–382.

HENIKOFF, S., 1987 Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155:** 156–157.

HERMAN, R. K., 1989 Mosaic analysis in the nematode *Caenorhabditis elegans*. J. Neurogenet. **5:** 1–24.

HORVITZ, H. R., 1988 Genetics of cell lineage, pp. 157–190 in *The Nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

JACOBSON, A., 1987 Purification and fractionation of poly(A)$^+$ RNA. Methods Enzymol. **152:** 254–261.

JOHNSON, C. D., J. B. RAND, R. K. HERMAN, B. D. STERN and R. L. RUSSELL, 1988 The acetylcholinesterase genes of *C. elegans*: identification of a third gene (*ace-3*) and mosaic mapping of a synthetic lethal phenotype. Neuron **1:** 165–173.

KOZAK, M., 1986 Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44:** 283–292.

KRAUSE, M., and D. HIRSH, 1987 A *trans*-spliced leader sequence on actin mRNA in Caenorhabditis elegans. Cell **49:** 735–761.

KYTE, J., and R. F. DOOLITTLE, 1982 A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:** 105–132.

LEVITT, A., and S. W. EMMONS, 1989 The Tc2 transposon in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **86:** 3232–3236.

LEWIS, J. A., J. S. ELMER, J. SKIMMING, S. MCLAFFERTY, J. FLEMING and T. MCGEE, 1987 Cholinergic receptor mutants of the nematode *Caenorhabditis elegans*. J. Neurosci. **7:** 3059–3071.

LEWIS, S. A., D. WANGA and N. J. COWAN, 1988 Microtubule-associated protein MAP2 shares a microtubule binding motif with Tau protein. Science **242:** 936–939.

MANIATIS, T., E. F. FRITSCH and SAMBROOK, 1982 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MCINTIRE, S. L., G. GARRIGA, J. WHITE, D. JACOBSON and H. R. HORVITZ, 1992 Genes necessary for directed axonal elongation or fasciculation in *C. elegans*. Neuron **8:** 307–322.

MELTON, D. A., P. A. KRIEG, M. R. REBAGLIATI, T. MANIATIS, K. ZINN and M. R. GREEN, 1984 Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12:** 7035–7056.

MIYATA, Y., M. HOSHI, E. NISHIDA, Y. MINAMI and H. SAKAI, 1986 Binding of microtubule-associated protein 2 and tau to the intermediate filaments reassembled from neurofilament 70-kDa subunit protein. J. Biol. Chem. **261:** 13026–13030.

NOBLE, M., S. A. LEWIS and J. COWAN, 1989 The microtubule binding domain of microtubule-associated protein MAP1B contains a repeated sequence motif unrelated to that of MAP2 and tau. J. Cell Biol. **109:** 3367–3376.

RAND, J. B., 1989 Genetic analysis of the *cha-1-unc-17* gene complex in Caenorhabditis. Genetics **122:** 73–80.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning*, Ed. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SCHWARTZ, D. C., and C. R. CANTOR, 1984 Separation of yeast chromosome-sized DNA by pulse field gradient gel electrophoresis. Cell **37:** 67–75.

SHEEN, J.-Y., and B. SEED, 1988 Electrolyte gradient gels for DNA sequencing. BioTechology **6:** 942–944.

SIDDIQUI, S. S., 1990 Mutations affecting axonal growth and guidance of motor neurons and mechanosensory neurons in the nematode *Caenorhabditis elegans*. Neurosci. Res. Suppl. **13:** 171–190.

SIDDIQUI, S. S., and J. G. CULOTTI, 1991 Examination of neurons in wild type and mutants of *Caenorhabditis elegans* using antibodies to horseradish peroxidase. J. Neurogenet. **7:** 193–211.

THOMAS, J. H., 1990 Genetic analysis of defecation in *Caenorhabditis elegans*. Genetics **103:** 43–64.

VIEIRA, J., and J. MESSING, 1987 Production of single-stranded plasmid DNA. Methods Enzymol. **153:** 3–11.

WATERSTON, R. H., 1988 Muscle, pp. 281–335 in *The Nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

WESSLER, S. R., 1991 The maize transposable *Ds1* element is alternatively spliced from exon sequences. Mol. Cell. Biol. **11:** 6192–6196.

WHITE, J. G., E. SOUTHGATE, J. N. THOMSON and S. BRENNER, 1986 The nervous system of *Caenorhabditis elegans*. Philos. Trans. R. Soc. Lond. Ser. B **314:** 1–340.

WICHE, G., C. OBERKANINS and A. HIMMLER, 1991 Molecular structure and function of microtubule-associated proteins. Int. Rev. Cytol. **124:** 217–273.

YUAN, J., M. FINNEY, N. TSUNG and H. R. HORVITZ, 1991 Tc4, a *Caenorhabditis elegans* transposable element with an unusual foldback structure. Proc. Natl. Acad. Sci. USA **88:** 3334–3338.

Communicating editor: M. T. FULLER