MOLECULAR AND CELLULAR BIOLOGY, Oct. 1989, p. 4119–4130
0270-7306/89/104119-12$02.00/0

Vol. 9, No. 10

Molecular Cloning and Sequencing of *ama-1*, the Gene Encoding the Largest Subunit of *Caenorhabditis elegans* RNA Polymerase II

DAVID M. BIRD[1,2]* AND DONALD L. RIDDLE[1]

Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211,[1] and Department of Nematology and Graduate Genetics Group, University of California Riverside, California 92521[2]

Received 15 November 1988/Accepted 21 June 1989

Two genomic sequences that share homology with *RpII215*, the gene encoding the largest subunit of RNA polymerase II in *Drosophila melanogaster*, have been isolated from the nematode *Caenorhabditis elegans*. One of these sequences was physically mapped on chromosome IV within a region deleted by the deficiency *mDf4*, 25 kilobases (kb) from the left deficiency breakpoint. This position corresponds to *ama-1* (resistance to α-amanitin), a gene shown previously to encode a subunit of RNA polymerase II. Northern (RNA) blotting and DNA sequencing revealed that *ama-1* spans 10 kb, is punctuated by 11 introns, and encodes a 5.9-kb mRNA. A cDNA clone was isolated and partially sequenced to confirm the 3′ end and several splice junctions. Analysis of the inferred 1,859-residue *ama-1* product showed considerable identity with the largest subunit of RNAP II from other organisms, including the presence of a zinc finger motif near the amino terminus, and a carboxyl-terminal domain of 42 tandemly reiterated heptamers with the consensus Tyr Ser Pro Thr Ser Pro Ser. The latter domain was found to be encoded by four exons. In addition, the sequence oriented *ama-1* transcription with respect to the genetic map. The second *C. elegans* sequence detected with the *Drosophila* probe, named *rpc-1*, was found to encode a 4.8-kb transcript and hybridized strongly to the gene encoding the largest subunit of RNA polymerase III from yeast, implicating *rpc-1* as encoding the analogous peptide in the nematode. By contrast with *ama-1*, *rpc-1* was not deleted by *mDf4* or larger deficiencies examined, indicating that these genes are no closer than 150 kb. Genes flanking *ama-1*, including two collagen genes, also have been identified.

The nematode *Caenorhabditis elegans* is being used by an increasing number of laboratories as a model organism for the molecular genetic analysis of metazoan development (33, 55). Understanding the control of gene expression in development will require a description of the RNA polymerases and how they interact with regulatory proteins and the DNA template. Much of our knowledge of RNA polymerase has come from studies of the procaryotic enzyme, in which biochemical and genetic methods have elucidated functions of the core polymerase ($\alpha_2\beta\beta'$) and its specificity factors (37, 56). Although eucaryotic nuclear RNA polymerases are structurally complex (48), recent molecular genetic approaches to understanding RNA polymerase II (RNAP II) have also proven useful (4–6, 39, 57). The objective of this work was to determine the structure of the gene encoding the largest subunit of this enzyme in *C. elegans* and to provide the basis for the molecular analysis of mutants affected in RNAP II function during development.

In most eucaryotes, RNAP II binds to α-amanitin, a toxin that inhibits RNA chain elongation (16). Mutations conferring resistance to α-amanitin have been isolated in a number of cell lines (13, 30, 49), in *Drosophila melanogaster* (26), and in *C. elegans* (44). Identification of transposon insertion mutants at a *Drosophila* amanitin resistance locus enabled the cloning of the *RpII215* gene (29, 47), which encodes the largest subunit of RNAP II (25). Probes spanning *RpII215* have subsequently been used to clone the analogous genes in yeast (31), mouse (17), and human (15) cells. Immunological approaches also has been used to clone genes encoding yeast RNA polymerase subunits, including the smaller ones (41, 51).

DNA sequence analysis of the yeast (3, 51), mouse (1), and *Drosophila* (32) genes has revealed regions of identity between the peptides from organisms widely separated by evolution, including the presence of a tandemly reiterated heptapeptide motif at the carboxyl terminus. In vivo analyses of deletion mutations have demonstrated that approximately half of this carboxyl-terminus domain (CTD) is required for viability in yeast (4, 39), *Drosophila* (57), and mouse (5) cells, although the entire CTD appears dispensable for accurate transcription in both *Drosophila* (57) and mammalian (34) in vitro-reconstituted systems.

The genetic analysis of *C. elegans* RNAP II is aimed at examining structure-function relationships for this enzyme in vivo. Mutants resistant to α-amanitin define two genes, *ama-1* and *ama-2* (42). RNAP II isolated from a strain homozygous for the *m118* allele of *ama-1* is 150 times less sensitive to inhibition by α-amanitin in vitro than is the wild-type enzyme (44). Twenty lethal *ama-1* alleles and several deficiencies and duplications covering *ama-1* have been generated (42, 43). The different lethal mutations cause developmental arrest at different times throughout the life cycle. Alleles causing complete loss of function result in arrest during the first larval stage, whereas certain other alleles result in sterility and are temperature sensitive. Fine-structure genetic mapping has ordered *m118* and 17 of the lethal alleles within the *ama-1* gene (12).

This paper describes the identification and cloning of two *C. elegans* genes, *ama-1* and *rpc-1*, homologous to the *Drosophila RpII215* locus. Using deficiency strains that carry deletions of *ama-1*, we first demonstrate that one of these genes, encoding a 5.9-kilobase (kb) mRNA, is located at a site corresponding to the genetic position of *ama-1*. Analysis of the cloned sequence confirms that *ama-1* en-

* Corresponding author.

codes the largest subunit of RNAP II and that this nematode peptide has features in common with the analogous peptides from other eucaryotes, including the CTD. Three of the four identified genes flanking *ama-1* encode collagens. One of these genes, located 11 kb to the left of *ama-1*, has been shown to be *dpy-13* (53), a frequently used genetic marker affecting body shape. The other gene detected with the *RpII215* probe encodes a 4.8-kb mRNA and also has extensive homology with the gene encoding the yeast RNAP III large subunit. Consequently, this gene is likely to encode an RNA polymerase subunit and has been named *rpc-1*. It has not yet been physically linked to *ama-1*.

MATERIALS AND METHODS

***C. elegans* strains and culture conditions.** Nematode stocks were grown on *Escherichia coli* OP50 (11), either on plates or in liquid S medium (50), and were harvested for nucleic acid isolation upon depletion of the food supply. After washing in M9 buffer (11), worms were incubated in M9 for 30 min at 20°C, with shaking, to clear any residual bacteria from the digestive tracts. Phenotypes of worms from the liquid cultures were confirmed by observation immediately before DNA or RNA isolation. The following strains were used: N2, wild type (11); DR680, *dpy-13(e184) ama-1(m118)/dpy-13(e184) ama-1(m118)* (43); *nT1/nT1* (23); and DR768, *mDf4/nT1[let(m435)]* (43).

RNA and DNA isolation. Nematode DNA was prepared as described by Emmons et al. (22), with minor modifications. The nematode pellet was not frozen and ground before proteinase K digestion; after the organic extractions, the solution was dialyzed against three changes of 10 mM Tris hydrochloride (pH 8.0)–1 mM EDTA. Subsequent centrifugation was found not to be necessary. RNA was isolated by ethanol precipitation from extracts of worms disrupted with a French pressure cell (82 MPa) in 6 M guanidinium hydrochloride essentially as described by Chirgwin et al. (14).

Plasmid and cosmid DNAs were isolated by the alkali-sodium dodecyl sulfate (SDS) method (10) and were purified by precipitation from 6.5% polyethylene glycol–400 mM NaCl for 1 h on ice. Bacteriophage DNA was extracted from CsCl-banded particles by phenol extraction and ethanol precipitation.

Detection of nematode sequences with a *Drosophila* probe. Plasmids p4.1 and p4.2, which span most of the *Drosophila RpII215* gene, were the gift of A. Greenleaf (47). *RpII215* encodes the largest subunit of RNAP II (9). Inserts purified from these plasmids were nick translated to high specific activity ($>10^8$ dpm/μg), pooled, and hybridized (54) to Southern blots of 5 μg of *C. elegans* genomic DNA digested with either *Hin*dIII, *Eco*RI, or *Bam*HI. Hybridization was done in 10% formamide–0.825 M Na^+ at 42°C for 16 h, followed by two brief washes at room temperature in 0.1% SDS–1.32 M Na^+ and two 30-min washes at 42°C in 0.1% SDS–330 mM Na^+.

Library screening. Phage DR#1 and DR#4 were isolated from a Charon 4 wild-type *C. elegans* (N2) library made from a partial *Eco*RI digest of genomic DNA (provided by T. Snutch and D. Baillie). The library was screened by the method of Benton and Davis (8) at reduced stringency (10% formamide, 0.825 M Na^+, 42°C hybridization; 330 mM Na^+, 0.1% SDS, 42°C wash) with probes p4.1 and p4.2. Phage from hybridizing plaques were isolated by three rounds of hybridization and plaque purification. Subclones from DR#1 were hybridized to an EMBL 4 library made from a partial *Sau*3A digest of N2 genomic DNA (gift of C. Link and W. B. Wood) to identify the overlapping clones DR#2 and DR#3.

Two arrays (contigs) of overlapping cosmid clones spanning the genomic lambda clones were kindly supplied by A. Coulson and J. Sulston. They identified the contigs by subjecting the lambda clones to a fingerprinting analysis and comparing restriction site patterns with those generated from cosmids (18).

Phage DB#2 was selected from a λgt10 cDNA library, constructed from poly(A)$^+$ RNA isolated from a synchronous population of *C. elegans* L2 larvae (provided by H.-M. Wang and M. Capecchi). Approximately 2×10^4 PFU was screened, at standard stringency (54), with a primer extension probe synthesized from DR#17 (a 3.7-kb *Eco*RI-*Hin*dIII fragment, generated from DR#1; see Fig. 7), and the single positive recombinant was subjected to three rounds of plaque purification.

Restriction analysis, subcloning, and electrophoresis. Recombinant and genomic DNA was digested, fractionated on agarose gels, blotted to nitrocellulose, and probed by standard methods (38). Restriction fragments were subcloned (after blunt ending with Klenow DNA polymerase where required) into the appropriate sites in either pUC118/119 or pGEM7Zf (Stratagene Corp.) and transformed into *E. coli* MC1061 or MV1193. Total RNA (10 μg per lane) was fractionated on agarose gels containing 10 mM CH_3HgOH and blotted onto nitrocellulose (52). Autoradiograms were quantified by densitometric scanning.

The actin gene probe was constructed from pRBA 4.8-2 (24), and the probe for the myosin heavy-chain gene (*unc-54*) was prepared from pRW28 (both provided by R. H. Waterston). The *col-1* probe was constructed from pJJ103, provided by J. Kramer (35).

DNA sequencing. Appropriate templates were generated from genomic and cDNA subclones, either by the random sonication method (21) or by using the *Exo*III-*Exo*VII nested-deletion approach. Single-stranded DNA was prepared after rescue with the helper phage M13K07 and subjected to dideoxy sequencing (46), using both universal and custom primers. Each genomic strand was sequenced at least twice, and appropriate constructs were exploited to sequence through all restriction sites used to make subclones. DB#2 subclones were sequenced to confirm the 3′ end of *ama-1* and also several inferred splice junctions. Formamide (25%) was included in some sequencing gels to enable resolution of compressions.

Nucleotide and inferred protein sequences were analyzed by using software from the University of Wisconsin Genetics Computer Group and Amersham Corp. (Staden-Plus).

RESULTS

Cloning of *C. elegans* sequences homologous to those encoding the large subunit of *Drosophila* RNAP II. Inserts from plasmids carrying the *Drosophila RpII215* gene were used as a probe to detect homologous *C. elegans* genomic sequences under low-stringency conditions (see Materials and Methods). Two *Eco*RI fragments (4.9 and 3.5 kb) were detected (Fig. 1), with the strongest hybridization to the larger fragment. Reductions in either hybridization or wash stringency failed to reveal additional bands discernible from nonspecific background hybridization. A *C. elegans* genomic library was then screened (see Materials and Methods) with the same low-stringency conditions used for the Southern blot shown in Fig. 1, and four plaques were selected. Restriction analysis of DNA from these clones revealed that they were two isolates each of two recombinants, DR#1 and DR#4. Hybridization of the *RpII215* probes to these clones showed

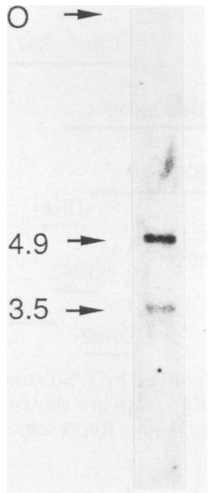

FIG. 1. Detection of *C. elegans* sequences homologous to the *Drosophila RpII215* locus. Genomic DNA (5 μg) from *C. elegans* was digested with *Eco*RI, fractionated on 1% agarose, blotted to nitrocellulose, and hybridized at low stringency with nick-translated *Drosophila* probes. The filter was subjected to two brief washes and autoradiographed. Transcript sizes (in kilobases) are indicated on the left. O, Origin.

that DR#1 contained the 4.9-kb *Eco*RI fragment detected by genomic blotting and that DR#4 contained the 3.5-kb *Eco*RI fragment (Fig. 2).

Chromosomal walking from DR#1 and DR#4. We sought additional clones to determine whether DR#1 and DR#4 were part of the same gene. Aliquots of DNA from the two lambda clones were sent to A. Coulson and J. Sulston (Medical Research Council, Cambridge, United Kingdom), who were able to identify a cosmid contig containing DR#4 sequences but not one overlapping DR#1 (see Materials and Methods). This result indicated that DR#1 and DR#4 are not contiguous sequences. Selected members of the contig spanning DR#4 are shown in Fig. 3A.

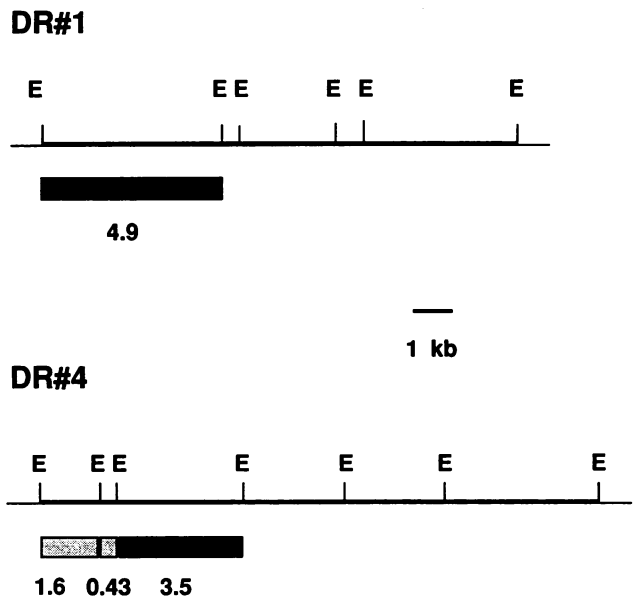

FIG. 2. Restriction maps of DR#1 and DR#4, showing positions of *Eco*RI sites (E) only. Solid boxes indicate fragments detected in Fig. 1. Stippled boxes identify fragments hybridizing to yeast RNAP III sequences as shown in Fig. 12.

To walk from DR#1, a 1.25-kb *Eco*RI-*Hin*dIII fragment from the left-most end of DR#1 was subcloned and used to screen an EMBL 4 genomic library of wild-type *C. elegans*. Seventeen positive plaques were selected and purified, and their DNAs were subjected to restriction analysis. Phage DR#3 was retained as the recombinant that extended most leftward from DR#1. However, since DR#3 added only approximately 3.8 kb to DR#1, a further library screening was performed, using a 1.55-kb *Eco*RI subclone derived from DR#3. Of 20 positive clones examined, 1 (DR#2) extended approximately 8.8 kb beyond DR#3. Both DR#2 and DR#3 were sent to Coulson and Sulston, who were then able to identify a 350-kb cosmid contig spanning these clones. Some of these cosmids are shown in Fig. 3B.

The fingerprints of the contigs containing DR#1 and DR#4 indicated that they did not overlap. To confirm this, colonies harboring cosmids spanning the entire DR#4 contig were grown on filters and probed with a range of subclones derived from the DR#1 contig, using the method of Grunstein and Hogness (27). In no case was hybridization between the two contigs observed (not shown). Since DR#1 and DR#4 were approximately in the center of their respective contigs, these sequences were judged to be a minimum of 250 kb apart. This indicated that DR#1 and DR#4 represent different genes.

DR#1 and DR#4 encode different transcripts. The 4.9-kb *Eco*RI fragment from DR#1 and the 3.5-kb *Eco*RI fragment from DR#4 were subcloned, generating DR#27 and DR#121, respectively. Inserts were isolated, nick translated, and used to probe Northern (RNA) filters of total *C. elegans* RNA. The DR#27 probe detected a 5.9-kb transcript, and the DR#121 probe detected a distinct transcript of 4.8 kb (Fig. 4). A 4.8-kb mRNA is too small to encode the largest subunit of *C. elegans* RNAP II, but a 5.9-kb transcript is sufficiently large (45).

DR#1 encodes *ama-1*. Genetic mapping experiments (43) placed *ama-1* on chromosome IV, 0.05 map unit to the right of *dpy-13* and close (<0.1 map unit) to the endpoint of the small-deficiency *mDf4*. This deletion of approximately 1.5 map units (0.5% of the genome) includes *ama-1* and *dpy-13* (53). Consequently, strains heterozygous for *mDf4* are hemizygous for *ama-1*, and this should be detectable as a dosage reduction by quantitative Southern blot analysis.

Genomic DNA was prepared from the *mDf4* heterozygote (*mDf4/nT1*) and from the two progenitor strains, *ama-1(m118)/ama-1(m118)* and the balancer *nT1/nT1*. A 5-μg sample of each was digested with *Eco*RI, fractionated on agarose, and blotted to nitrocellulose. Filters were probed with nick-translated inserts from DR#27 (4.9 kb) or DR#121 (3.5 kb) plus either an actin or myosin heavy-chain gene probe. Neither of the four actin genes nor the myosin heavy-chain gene is within *mDf4*, so these probes served to normalize the amount of DNA per lane. The result obtained by using the DR#27 probe is shown in Fig. 5A. The autoradiogram was traced with a densitometer (Fig. 5B), and the area under the peaks was integrated. Relative to the controls, the amount of hybridization to the *mDf4* heterozygote was reduced to 0.40.

The hybridization result indicated that the 4.9-kb fragment from DR#1 is within *mDf4* and therefore may encode *ama-1*. The strain tested for DR#1 copy number is a balanced heterozygote carrying the recessive-lethal *mDf4* over a recessive-lethal derivative of *nT1*. A possible reason why the

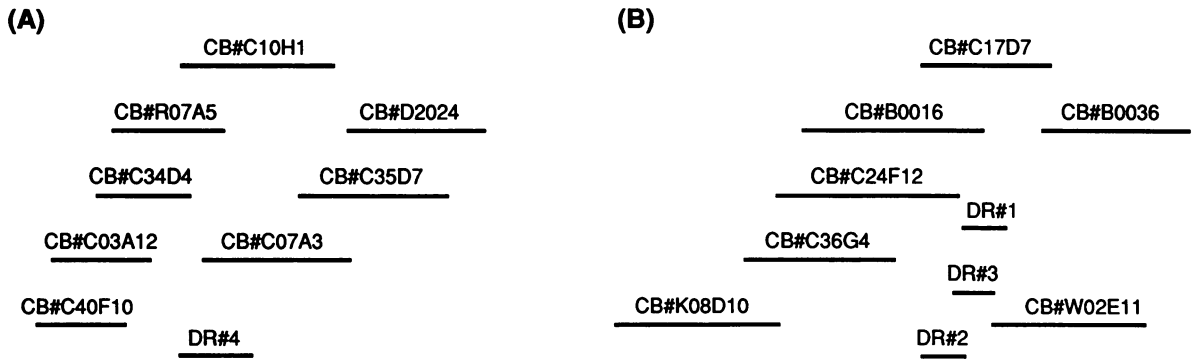

FIG. 3. Arrays of overlapping cosmid clones, or contigs, that span DR#4 (A) or DR#1 (B), identified by J. Sulston and A. Coulson by use of a fingerprinting technique (18). Lambda clones DR#2 and DR#3, isolated by walking from DR#1, also are shown (B). The length of each line is proportional to the number of *Hin*dIII sites in each molecule (on average, each cosmid is 34 kb). Representative cosmids from each contig are shown.

reduction in band intensity was 60% rather than 50% is that *mDf4* homozygotes accumulate preferentially in the population as arrested embryos.

When the DR#121 insert was used as the probe, no difference in hybridization intensity to the 3.5-kb genomic fragment was observed between either of the progenitor or the *mDf4* heterozygous strains relative to the myosin control fragments (not shown). Consequently, the 3.5-kb *Eco*RI fragment is not within *mDf4*, eliminating the possibility that DR#4 encodes *ama-1*. These blot experiments were repeated by using DNA isolated from strains heterozygous for the larger deficiencies (approximately 1% of the genome), *mDf5* and *mDf9*, both of which delete *ama-1* (43). In each case (not shown), sequences from DR#1 were found to be spanned by the deletions, and sequences from DR#4 were found not to be. This result placed DR#1 and DR#4 at least 0.5 map unit apart (approximately 150 kb) and is consistent with the result obtained by probing the cosmid contig.

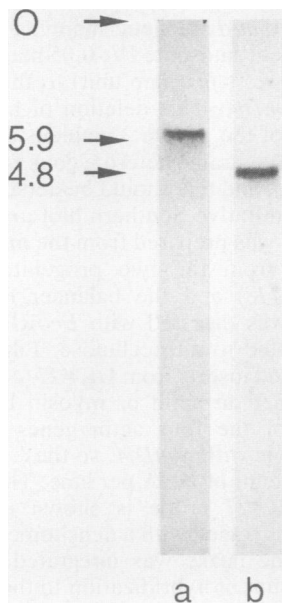

FIG. 4. Northern blots showing that DR#1 and DR#4 encode different transcripts. Total *C. elegans* RNA (10 μg) was fractionated on a 1% agarose gel containing 10 mM CH_3H_gOH, blotted to nitrocellulose, and probed with either DR#27 (constructed from DR#1) (a) or DR#121 (constructed from DR#4) (b). Transcript sizes (in kilobases) are indicated on the left. O, Origin.

The *mDf4* breakpoint. To more precisely correlate DR#1 with the genetic map, the physical location of the *mDf4* breakpoint was identified. Since *ama-1* is closely linked genetically (<0.1 map unit) to the left breakpoint of *mDf4*, it seemed likely that one of the cosmids from the DR#1 contig might span this site. Whole cosmids from this contig were nick translated and probed to Southern blots of *Eco*RI digests of N2 (wild type) and DR768 (*mDf4/nT1*) DNA (not shown). Although the presence of repeated elements com-

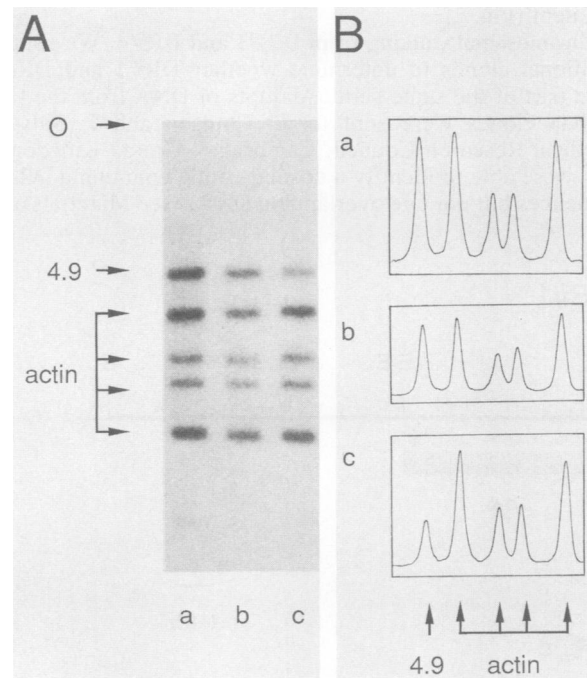

FIG. 5. Demonstration that DR#1 sequences are within the region deleted by *mDf4*. Genomic DNA (5 μg) from a balanced strain heterozygous for *mDf4* and from its progenitors, an *ama-1* (*m118*) homozygote and a strain homozygous for the balancer chromosome *nT1*, were digested with *Eco*RI, fractionated on 1% agarose, and blotted to nitrocellulose. (A) Autoradiogram of a Southern blot of *Eco*RI digests of *nT1/nT1* (two gene copies; lane a), *ama-1/ama-1* (two gene copies; lane b), and *mDf4/nT1* (one gene copy; lane c) hybridized with an actin gene probe plus DR#27 insert. Actin gene bands are indicated; sizes are in kilobases. O, Origin. (B) Densitometric trace of the autoradiogram in panel A; labeling is the same.

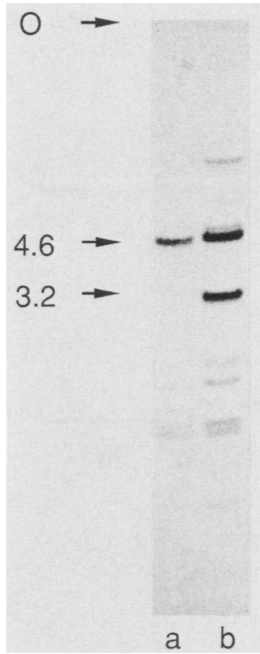

FIG. 6. Mapping the *mDf4* breakpoint. Approximately 5 μg of genomic DNA from N2 (wild type) (a) or DR768 (*mDf4/nT1*) (b) worms were digested with *Eco*RI, fractionated on 1% agarose, blotted to nitrocellulose, and probed with DR#8 (a 4.6-kb *Eco*RI fragment 24.5 kb to the left of *ama-1*). Sizes are in kilobases. O, Origin.

plicated the interpretation, probing the DR768 DNA with CB#B0016 (Fig. 3B) appeared to detect a new band at 3.2 kb, with concomitant reduction of a band at 4.6 kb. To verify this observation, the 4.6-kb *Eco*RI fragment was subcloned, generating DR#8, and the insert was used as a probe. The result confirmed that the *mDf4* breakpoint lies within the 4.6-kb *Eco*RI fragment (Fig. 6). Additional genomic blots, with DR768 and N2 DNA cut with a range of enzymes, positioned the breakpoint close to the left end of this fragment (data not shown).

Identification of the *mDf4* breakpoint close to DR#1 provided strong corroborating evidence that DR#1 encodes *ama-1*, and the result allowed the cosmid contig to be oriented with respect to the genetic map. Restriction mapping of CB#B0016 (53) positioned the left end of the 4.6-kb fragment 24.5 kb from DR#16 (Fig. 7 and 8), a distance consistent with the genetic distance of *ama-1* from the left breakpoint of *mDf4* (on average, 0.1 map unit = 30 kb).

Transcript mapping *ama-1*. To delineate the approximate boundaries of *ama-1*, subclones spanning 23 kb around the 4.9-kb *Eco*RI fragment were constructed and used as probes to Northern blots of total *C. elegans* RNA. Hybridizations initially were performed by using the inserts from DR#27, DR#29, DR#35, DR#36, and DR#43 (identified in Fig. 7). This experiment served to localize *ama-1* within a region of approximately 11.5 kb. None of the subclones to the left of the DR#36 (3.5-kb) probe detected the *ama-1* transcript, suggesting that one terminus of this gene might lie within or close to the boundary of this fragment, possibly within DR#16. The DR#16 probe strongly detected the 5.9-kb mRNA. Similarly, no hybridization to the right of the DR#27 probe was observed, placing the other terminus of *ama-1* in this fragment. Attempts to more precisely locate this terminus were hampered by the paucity of restriction sites in the right 75% of this molecule. No sites were identified for the 14 five-base or six-base cutters tested, and of the four-base cutters examined, only *Tha*I was found to be of use. The 2.15-kb *Tha*I-*Eco*RI molecule (DR#30) was the most rightward fragment identified that hybridized to the *ama-1* message. These Northern blots established the minimum size of the *ama-1* transcribed region as 7.3 kb, suggesting the presence of at least 1.4 kb of introns.

Probes DR#28, DR#43, and DR#37 detected transcripts other than *ama-1* (Fig. 7). DR#43 detected a 750-base-pair (bp) RNA in approximately fivefold lower abundance than *ama-1*, and it is yet to be identified. DR#28 and DR#37 both detected smears of very abundant RNAs between 1.2 and 1.4 kb. Reprobing these same filters with pRW28 (myosin heavy chain) yielded a discrete band at 6 kb (not shown), indicating that the smears were not the result of RNA degradation. Similar smears have been obtained by using collagen gene probes (36). To test whether DR#28 and DR#37 might encode members of the collagen multigene family, *Eco*RI digests of DR#1 and CB#B0016 were challenged with a *C. elegans col-1* probe. In each case, the collagen probe hybridized to the *Eco*RI fragments used to generate DR#28 and DR#37. These two collagen genes have been assigned the names *col-33* and *col-34*, respectively. A collagen gene bank was screened with DR#32 (J. Kramer, personal communication), which carries single-copy DNA flanking *col-33* (Fig. 8). This screening revealed that *col-33* is the same collagen gene carried by CH#λCG41 (J. Kramer, unpublished data). Further analysis of CH#λCG41 indicated that this clone overlaps DR#1 and extends 1.6 kb further to the right. A similar screen was not performed for *col-34*. The positions of *ama-1* and adjacent genes are summarized in Fig. 8.

Nucleotide sequence of *ama-1*. A genomic sequence of 10,019 bp spanning *ama-1*, from 99 bp to the left of DR#16 to 1,639 bp to the right of DR#31 (Fig. 7), is shown in Fig. 9. In addition, 1,356 bp of DNA sequence has been determined from the cDNA clone DB#2 (underlined in Fig. 9). Since the sequence from DB#2 includes a stretch of the poly(A) tract, the most 3′ transcribed residue in the mature *ama-1* message can unambiguously be assigned to position 9979, 21 bases downstream of an AATAAA motif.

Statistical analyses of codon distribution of the genomic sequence, including plots of negentropy and local deviation in base composition, implicated the presence of 11 introns. Examination of the borders of these presumptive introns revealed that they all are flanked by *C. elegans* donor and acceptor consensus splice sequences (55). There is an additional acceptor consensus at position −5 (Fig. 9). However, examination of sequences upstream (not shown) suggests that the closest donor consensus is more than 500 bp upstream; that *ama-1* extends beyond that point is inconsistent with the transcript mapping results (Fig. 7). Alignment of the cDNA and genomic sequences confirmed the theoretical delineation of three introns, which span a total of 1,619 bp. Removal of all 11 inferred introns leaves a sequence of 6,009 bp. An ATG codon, which begins an open reading frame that links the 12 exons, occurs 40 bp from the start of this sequence.

Alignment of the inferred *ama-1* product with other RNA polymerases (see below) strongly suggests that this assignment of the translation initiator is correct. The exons encode peptides of 28, 37, 109, 237, 102, 260, 148, 244, 530, 69, 62, and 33 amino acids, respectively. The resulting 1,859-residue protein has a calculated molecular weight of 204,630, which is in close agreement with the 200,000 molecular weight of the II_a subunit of nematode RNAP II, estimated from

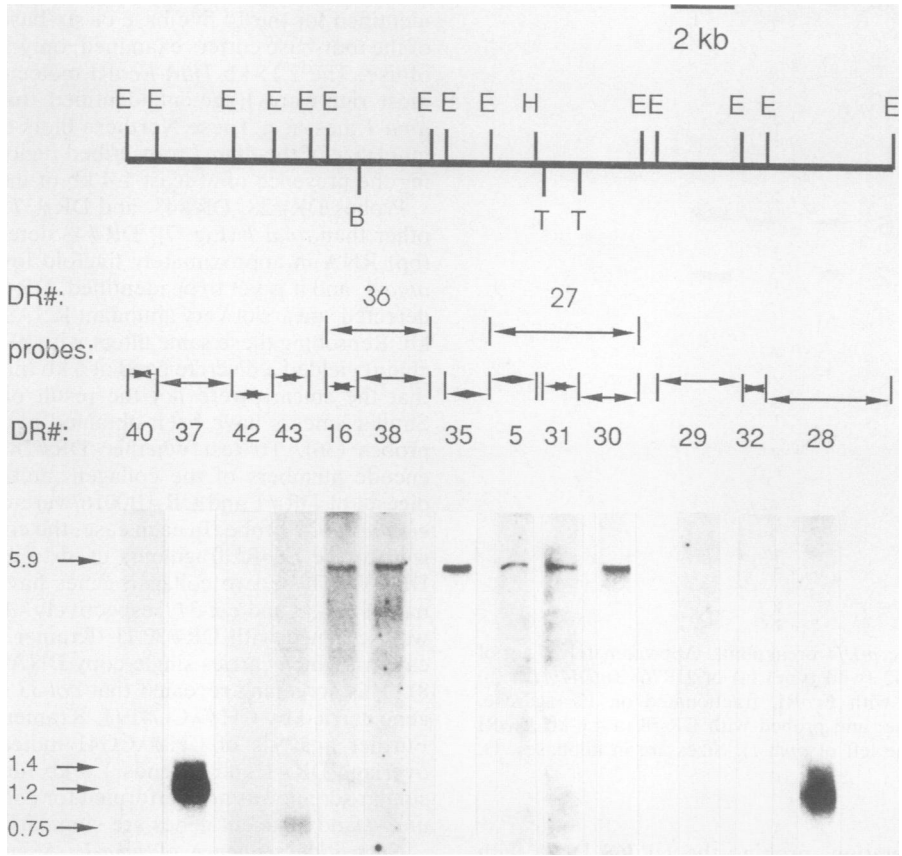

FIG. 7. Mapping transcripts in the region around *ama-1*. The overlapping lambda clones DR#1, DR#2, and DR#3 and the overlapping cosmid CB#B0016 (see Fig. 3B) were restriction mapped; only the sites used to construct the named subclones are shown (E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI; T, *Tha*I). Nick-translated probes were generated from these subclones as indicated and hybridized to Northern filters of 10 μg of total *C. elegans* RNA from an asynchronous culture. Filters probed with DR#28 or DR#37 were autoradiographed for 30 min. All other filters were exposed for 16 h with an intensifying screen. Additional periods of exposure (including the DR#28 and DR#37 filters) failed to show additional bands of specific hybridization. However, longer exposure did reveal bands corresponding to rRNA (most evident in the DR#43 lane). Some filters were from different gels and have been arranged so that bands of the same size (in kilobases) are aligned.

SDS-polyacrylamide gel electrophoresis (44). The length of neither the 5′ untranslated region nor the poly(A) tail has yet been determined, but the distance from the ATG to the start of the poly(A) tail (5,970 nucleotides) is in agreement with the measured size of the mature mRNA (5.9 kb). In addition, the DNA sequence orients the gene with its 5′ end to the left on the genetic map.

Amino acid homologies between the *ama-1* product and

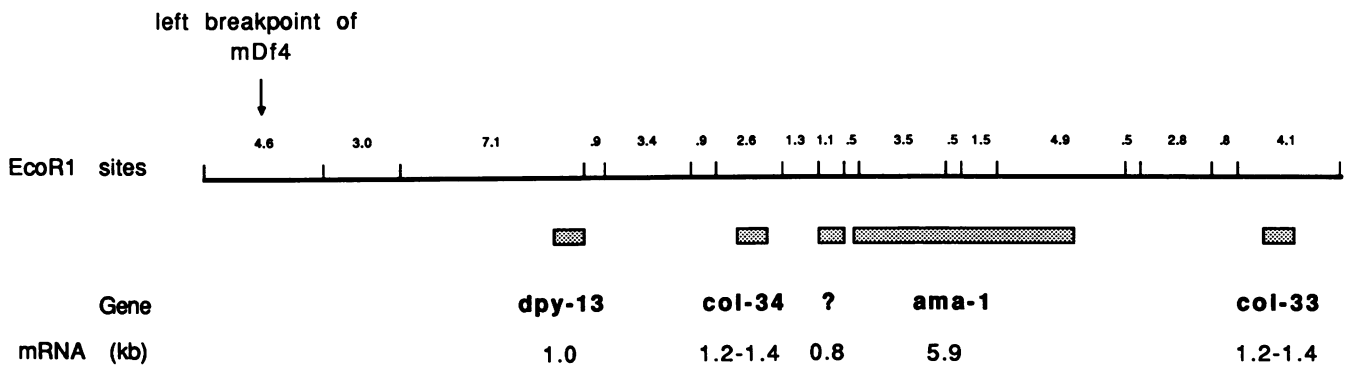

FIG. 8. Physical map of the *dpy-13–ama-1* region on chromosome IV, adapted from von Mende et al. (53). The region was mapped by using single and double restriction digests of cosmid CB#B0016 and three overlapping lambda clones, using the endonucleases *Eco*RI, *Eco*RV, *Bam*HI, *Hin*dIII, and *Sal*I (only *Eco*RI sites are shown). All *Eco*RI fragments except the 3.4-kb fragment were subcloned and hybridized to Southern blots of the mapping digests to ensure correct positions. The direction of transcription of *dpy-13* and *ama-1* is from left to right. The exact sizes of the collagen genes, *col-33* and *col-34*, and their exact positions between flanking *Eco*RI sites are not known.

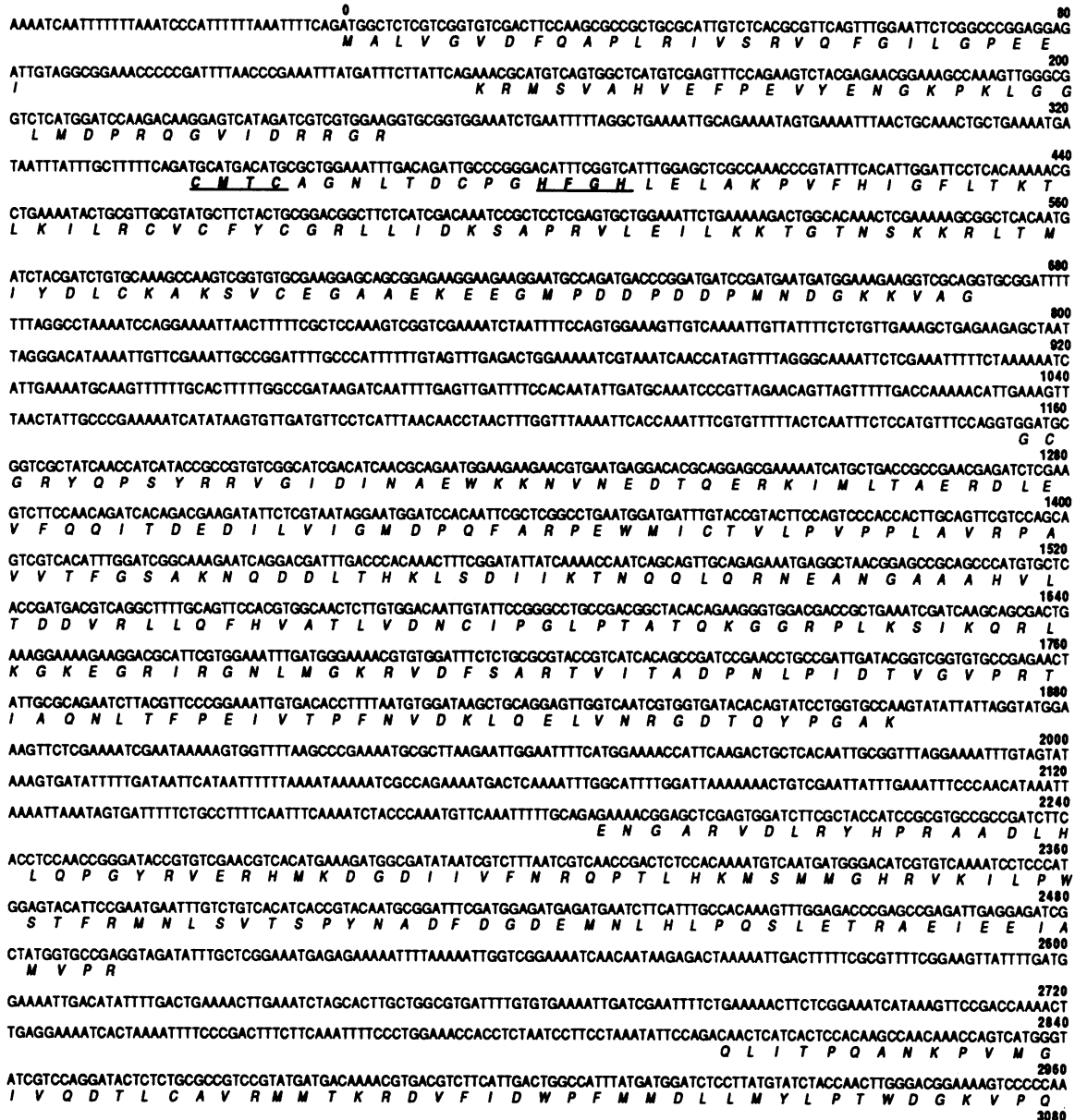

FIG. 9. Genomic, partial cDNA, and deduced amino acid sequences of *ama-1*. The genomic DNA sequence from 99 bp to the left of DB#16 to 1,639 bp to the right of DB#31 (Fig. 7) is shown. The first base of the presumed initiator codon has been assigned position 0. Sequences that additionally were derived from the cDNA clone DB#2 are underlined. Position 9979 is the last residue before the poly(A) tail; 21 nucleotides upstream of this position is the likely AATAAA processing signal (open letters). The inferred amino acid sequence of subunit II_a is shown in one-letter code below the DNA sequence. The zinc finger consensus sequence, beginning at amino acid 66, is underlined. Assignment of exon borders is discussed in the text.

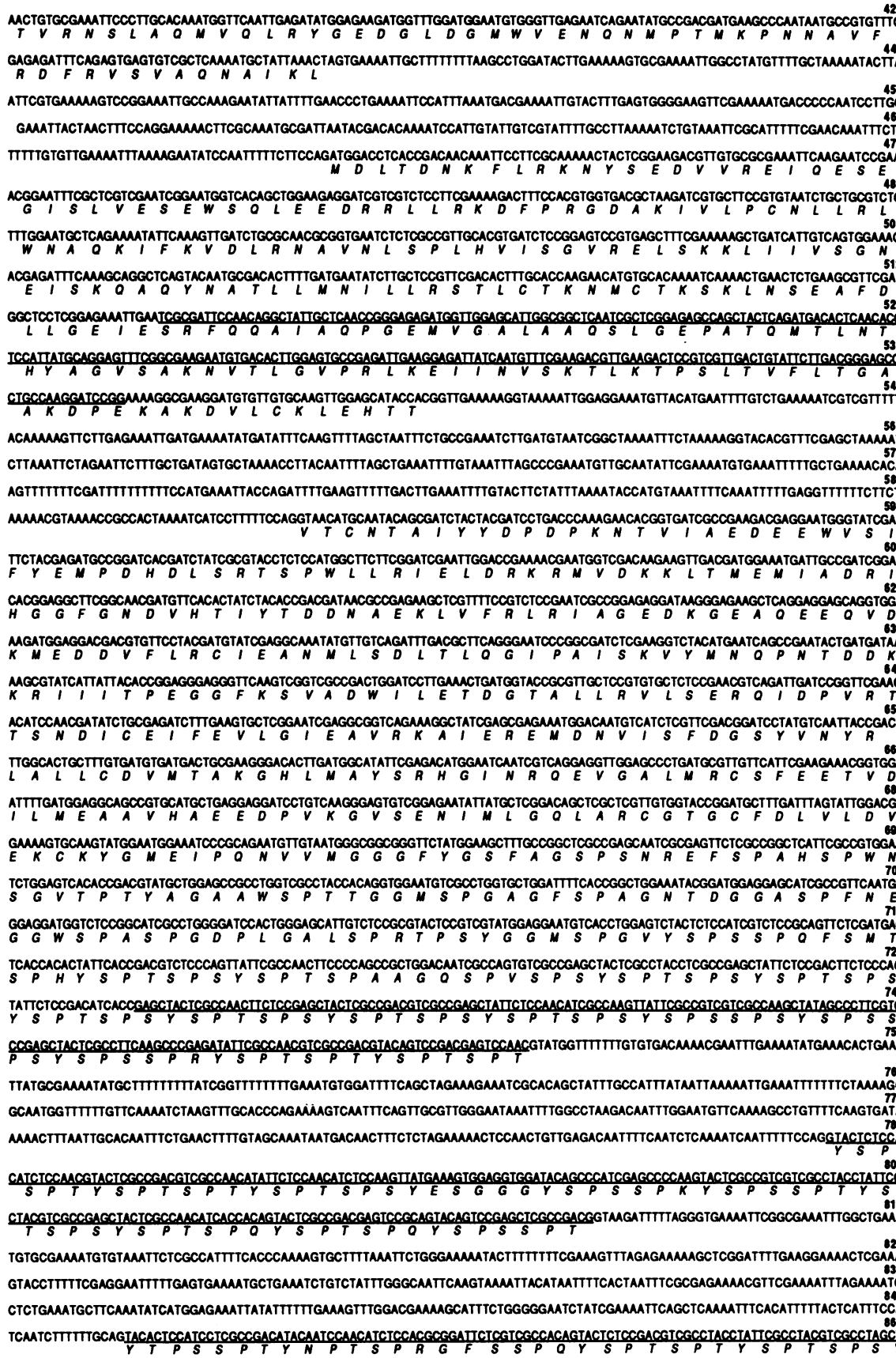

Fig. 9—*Continued*

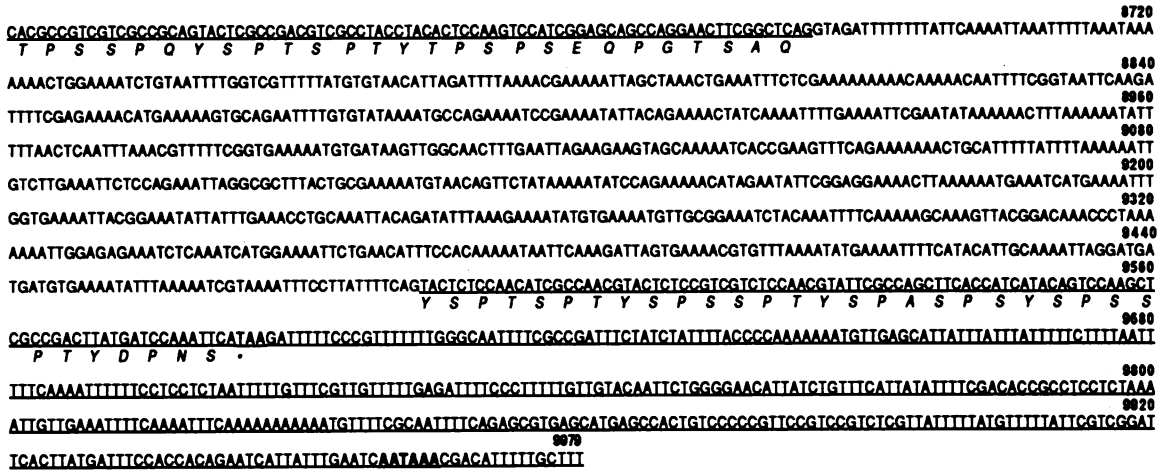

Fig. 9—*Continued*

RNA polymerase subunits from other organisms. The sequence of the inferred *ama-1* product (Fig. 9) shares substantial identity with the sequences of the largest RNAP II subunits from other organisms, including yeasts (3, 51), *D. melanogaster* (32), and mice (1). This finding confirms the identity of *ama-1* as encoding the RNAP II_a protein in *C. elegans*.

The first 1,500 residues of the *ama-1* protein share seven extended blocks of high identity with many RNA polymerase subunits, including those from viral and procaryotic enzymes. These regions have been discussed in detail elsewhere (1, 3, 32). In addition to these specific domains, there is extensive overall identity between the worm II_a subunit and the analogous proteins from other eucaryotes. For example, by using the Gap program, similarity between the *ama-1* and mouse proteins was found to be greater than 80%. Figure 10 shows dot matrix comparisons between the *ama-1* product and the analogous peptides from yeast, *Drosophila*, and mouse cells, respectively. A perfect match of five consecutive residues was required for a dot to be plotted. Two features are apparent. Firstly, the initial 1,500 amino acids of the *C. elegans* polypeptide share much greater homology with the metazoan proteins than with the yeast subunit; homology with the yeast sequence is restricted primarily to the seven high-homology blocks (1, 3, 32). A single zinc finger domain begins at amino acid 66.

Second, the carboxyl terminus of the *ama-1* product is highly homologous with the CTDs in the other three proteins. This identity appears uniformly high throughout the extent of the yeast repeat. In comparison, the homology between the nematode and fly domains fluctuates along its length, with the highest identity at approximately the center of the *Drosophila* domain. The first half of the mouse repeat is more similar to the nematode repeat than is the second half. The CTD of the *ama-1* product is composed of a tandemly reiterated heptamer with a consensus Tyr Ser Pro Thr Ser Pro Ser, although the seventh position is quite variable, especially beyond repeat 23 (Fig. 11). As is the case with *D. melanogaster* (32), the heptamer is reiterated 42 times; the yeast genome has 26 repeats (3), and the mouse genome has 52 (1). By contrast with these species, the CTD of the *C. elegans* subunit is not terminated by a short acidic sequence.

DR#4 may encode the large subunit of RNAP III. In addition to detecting *ama-1*, the *Drosophila RpII215* probe also hybridized to *C. elegans* sequences spanned by DR#4. Using the same fly probe, Ingles et al. (31) have detected three regions in the yeast genome and demonstrated that two of these regions encode the large subunits of RNAP II and III, respectively (3). By analogy with yeast cells, it seemed possible that DR#4 might span the gene encoding the *C. elegans* RNAP III large subunit (4.8 kb of transcript is sufficient to encode this 140-kilodalton peptide). Evidence consistent with this notion was generated by probing a Southern filter of an *Eco*RI digest of DR#4 with sequences spanning the large subunit of yeast RNAP III (2.6- and 2.3-kb *Eco*RI fragments derived from RPO31, provided by C. J. Ingles). At reduced stringency (see Materials and Methods), the RPO31 probe hybridized to three *Eco*RI bands in DR#4; the 3.5-kb fragment detected with the

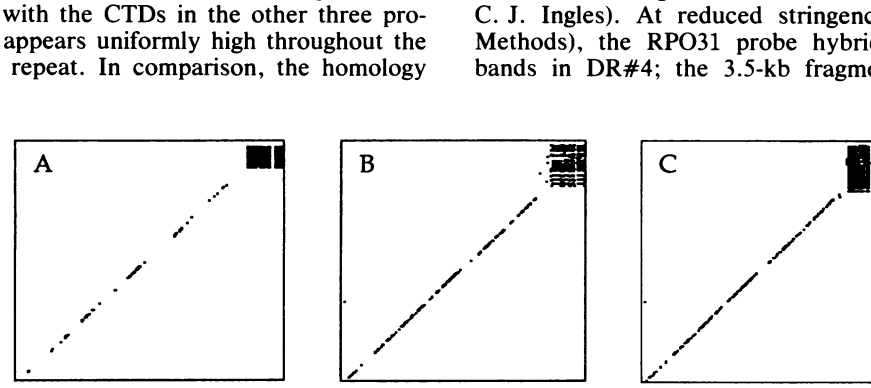

FIG. 10. Matrix comparing the deduced *C. elegans* RNAP II_a amino acid sequence (horizontal axes) with the inferred sequence of the II_a subunits of the yeast (A), *Drosophila* (B), and mouse (C) genomes (vertical axes). Using the program Compare, run in word comparison mode, it was specified that a match of five consecutive identical amino acids between the sequences was required to generate a point. Data were plotted by using the program DotPlot.

1	L	S	P	R	T	P	S	
2	Y	G	G	M	S	P	G	V
3	Y	S	P	S	S	P	Q	
4	F	S	M	T	S	P	H	
5	Y	S	P	T	S	P	S	
6	Y	S	P	T	S	P	A	A G
7	Q	S	P	V	S	P	S	
8	Y	S	P	T	S	P	S	
9	Y	S	P	T	S	P	S	
10	Y	S	P	T	S	P	S	
11	Y	S	P	T	S	P	S	
12	Y	S	P	T	S	P	S	
13	Y	S	P	T	S	P	S	
14	Y	S	P	S	S	P	S	
15	Y	S	P	S	S	P	S	
16	Y	S	P	S	S	P	R	
17	Y	S	P	T	S	P	T	
18	Y	S	P	T	S	P	T	
19	Y	S	P	T	S	P	T	
20	Y	S	P	T	S	P	T	
21	Y	S	P	T	S	P	S	
22	Y	E	S	G	G	G		
23	Y	S	P	S	S	P	K	
24	Y	S	P	S	S	P	T	
25	Y	S	P	T	S	P	S	
26	Y	S	P	T	S	P	Q	
27	Y	S	P	T	S	P	Q	
28	Y	S	P	S	S	P	T	
29	Y	T	P	S	S	P	T	
30	Y	N	P	T	S	P	R	
31	G	F		S	S	P	Q	
32	Y	S	P	T	S	P	T	
33	Y	S	P	T	S	P	S	
34	Y	T	P	S	S	P	Q	
35	Y	S	P	T	S	P	T	
36	Y	T	P	S	P	S	E	
37	Q	P	G	T	S	A	Q	
38	Y	S	P	T	S	P	T	
39	Y	S	P	S	S	P	T	
40	Y	S	P	A	S	P	S	
41	Y	S	P	S	S	P	T	
42	Y	D	P	N	S	*		
	Y	S	P	T/S	S	P	S/T	CONSENSUS

FIG. 11. The CTD. The deduced amino acid sequence of the *ama-1* product, from residue 1567 to the carboxyl terminus, is shown. Numbering corresponds to that used for mouse (1) and *D. melanogaster* (32). A consensus sequence of the heptamer is shown below the sequence.

Drosophila RpII215 probe and two additional molecules of 1.6 and 0.43 kb (Fig. 12). These three *Eco*RI fragments are contiguous in the genome (Fig. 2). No hybridization of the yeast RPO31 probes to DR#1 was observed, although some weak, nonspecific background was detected.

These hybridization results suggest that DR#4 encodes an RNA polymerase subunit or a closely related protein. Although it seems likely that this protein is the largest subunit of RNAP III, additional confirming data are required. Consequently, the gene has been named *rpc-1* (RNA polymerase C), which is both in accord with *C. elegans* nomenclature rules (28) and consistent with RNA polymerase nomenclature conventions.

DISCUSSION

We have cloned two unique regions from the *C. elegans* genome that share homology with sequences encoding the large subunit of *Drosophila* RNAP II. One of these clones, DR#1, was found to be derived from sequences that are deleted by *mDf4*, a small deletion that covers *ama-1* (a gene previously shown to encode an RNAP II subunit). We identified the left breakpoint of this deficiency by performing a series of hybridizations to digests of genomic DNA from a strain heterozygous for *mDf4* with probes extending from DR#1. The physical distance from this breakpoint to DR#1 corresponds to its genetic distance from *ama-1*, providing strong evidence that DR#1 encodes *ama-1*. The identification and cloning of the *dpy-13* gene between *ama-1* and the *mDf4* breakpoint (53) further confirmed this conclusion.

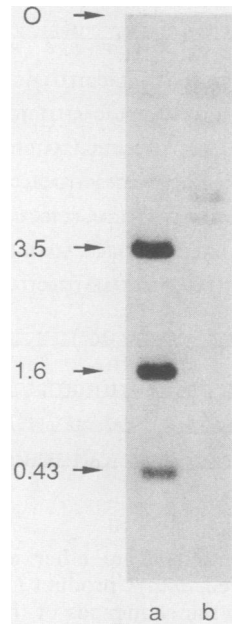

FIG. 12. Southern blot of *Eco*RI digests of DR#4 (lane a) and DR#1 (lane b), probed at low stringency (see Materials and Methods) with sequences spanning the largest subunit of yeast RNAP III. Sizes are shown in kilobases. O, Origin.

Northern blotting showed that *ama-1* encodes a 5.9-kb transcript, which is sufficiently large to encode the largest subunit of *C. elegans* RNAP II (45). Elucidation of the entire genomic DNA sequence confirms that this is indeed the *ama-1* product. The inferred protein has substantial identity with its counterparts in mouse and *Drosophila* cells. In addition, there are seven regions that share significant homology with many RNA polymerases, including procaryotic enzymes. The likely functions of these domains and specific motifs within them, including the zinc finger homology at residues 66 to 72, have been discussed elsewhere (1, 3, 32). In common with other eucaryotes, the largest subunit of *C. elegans* RNAP II has a CTD composed of tandem reiterations of the motif Tyr Ser Pro Thr Ser Pro Ser. The nematode protein differs slightly from that of other organisms in that it lacks a short acidic extension from the CTD. Potential roles for the CTD in transcription have been proposed and extensively discussed (4, 6, 32, 34, 39, 57).

Although sequences upstream of the ATG remain to be characterized, it appears that *ama-1* is broken into 12 exons. On the basis of amino acid sequence alignments, Jokerst et al. (32) have noted a correspondence between the location of the three introns in *Drosophila RpII215* and three of the 27 introns in the analogous mouse gene. Alignment of *ama-1* with these two genes reveals that the position of the first intron is common to all three species and that the 6th intron is equivalent to the 2nd in *D. melanogaster* and the 14th in mouse cells. By contrast with other organisms examined, the nematode CTD is not encoded by a single exon.

In addition to confirming that *ama-1* encodes the largest RNAP II subunit, the DNA sequence orients the gene with its 5′ end to the left on the genetic map. The genetic fine-structure map of *ama-1* (12) can now be aligned with the DNA sequence of the gene. The order of mutations within the gene will serve as a guide to the sequence analysis of the mutant alleles, with the aim of understanding enzyme structure-function relationships.

In yeast cells, the *Drosophila RpII215* probe detects two sequences in addition to those encoding the large subunit of RNAP II (31), and one of these encodes the RNAP III large subunit (3). An RNAP III probe from yeast cells strongly hybridized to *rpc-1* but not to *ama-1*. Furthermore, the size of the *rpc-1* transcript (4.8 kb) is entirely consistent with this gene encoding a 140-kilodalton polypeptide. Characterization of cosmid contigs around *ama-1* and *rpc-1* indicated that these clones are not contiguous. Use of the *mDf5* and *mDf9* deficiency strains confirmed this and placed the putative RNAP III large subunit gene at least 0.5 map units from *ama-1*. Cosmids spanning both *ama-1* and *rpc-1* have been hybridized in situ (2) to spreads of *C. elegans* chromosomes, and the data placed both sequences on chromosome IV (D. Albertson, personal communication). Recently, the two contigs have been joined with others by hybridization to *C. elegans* inserts in yeast artificial chromosomes (19), but they have not yet been joined to each other. Thus, *ama-1* and *rpc-1* are judged to be more than 600 kb apart (A. Coulson, J. Sulston, and R. H. Waterston, personal communication).

In addition to isolating and characterizing *ama-1* and *rpc-1*, we also have identified three additional genes in the immediate vicinity of *ama-1*, two flanking collagen genes (*col-33* and *col-34*) and an as yet unidentified gene encoding a 750-bp transcript. A third collagen gene, *dpy-13* (short, dumpy body shape), lies approximately 6 kb to the left of *col-34* (53). Hence, *ama-1* appears to be nested in a cluster of collagen genes. The *C. elegans* collagen multigene family consists of 50 to 150 members (20), most of which are thought to encode structural components of the cuticle (40). These genes are typically expressed in a stage-specific manner at the larval molts (40), and they are expressed specifically in the hypodermis, the tissue that secretes the cuticle components. By contrast, *ama-1* performs a housekeeping function in all cells, and its expression is unlikely to be limited to larval molts. It also will be interesting to determine whether the transcripts of this housekeeping gene are *trans* spliced with the *C. elegans* 22-nucleotide 5′ leader sequence (7). It is possible that the potential splice acceptor sequence adjacent to the ATG is a *trans*-splicing acceptor. This would be consistent with the data indicating that the 5′ untranslated region must be short.

The gene encoding the 750-bp transcript is between *col-34* and *ama-1*, very close to *ama-1* on its left. Sequence analysis of this gene might reveal whether it is related to collagen function or to RNA polymerase function, and hybridization studies could reveal if it is developmentally regulated.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM36674 from the National Institutes of Health.

We thank T. Snutch and D. Baillie, H.-M. Wang, and M. Capecchi, and W. B. Wood for the gifts of *C. elegans* libraries, A. Greenleaf for plasmids p4.1 and p4.2, J. Ingles for RPO31 clones, R. Waterston for plasmids pRBA 4.8-2 and pRW28, J. Kramer for pJJ103, and A. Coulson and J. Sulston for cosmids. Some DNA sequencing was performed by M. A. Wilson.

LITERATURE CITED

1. **Ahearn, J. M., Jr., M. S. Bartolomei, M. L. West, L. J. Cisek, and J. L. Corden.** 1987. Cloning and sequence analysis of the mouse genomic locus encoding the largest subunit of RNA polymerase II. J. Biol. Chem. **262**:10695–10705.
2. **Albertson, D. G.** 1985. Mapping muscle protein genes by *in situ* hybridization using biotin-labeled probes. EMBO J. **4**:2493–2498.
3. **Allison, L. A., M. Moyle, M. Shales, and C. J. Ingles.** 1985. Extensive homology amongst the large subunits of eukaryotic and prokaryotic RNA polymerases. Cell **42**:599–610.
4. **Allison, L. A., J. K.-C. Wong, V. D. Fitzpatrick, M. Moyle, and C. J. Ingles.** 1988. The C-terminal domain of the largest subunit of RNA polymerase II from *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals: a conserved structure with an essential function. Mol. Cell. Biol. **8**:321–329.
5. **Bartolomei, M. S., and J. L. Corden.** 1987. Localization of an α-amanitin resistance mutation in the gene encoding the largest subunit of mouse RNA polymerase II. Mol. Cell. Biol. **7**: 586–594.
6. **Bartolomei, M. S., N. F. Halden, C. R. Cullen, and J. L. Corden.** 1988. Genetic analysis of the repetitive carboxy-terminal domain of the largest subunit of mouse RNA polymerase II. Mol. Cell. Biol. **8**:330–339.
7. **Bektesh, S., K. V. Doren, and D. Hirsh.** 1988. Presence of the *Caenorhabditis elegans* spliced leader on different mRNAs and in different genera of nematodes. Genes. Dev. **2**:1277–1283.
8. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science **196**:180–182.
9. **Biggs, J., L. L. Searles, and A. L. Greenleaf.** 1985. Structure of the eukaryotic transcription apparatus: features of the gene for the largest subunit of *Drosophila* RNA polymerase II. Cell **42**:611–621.
10. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
11. **Brenner, S.** 1974. The genetics of *C. elegans*. Genetics **77**:71–94.
12. **Bullerjahn, A. M. E., and D. L. Riddle.** 1988. Fine-structure genetics of *ama-1*, an essential gene encoding the large subunit of RNA polymerase II in *Caenorhabditis elegans*. Genetics **120**:423–434.
13. **Chan, V. L., G. F. Whitmore, and L. Siminovitch.** 1972. Mammalian cells with altered forms of RNA polymerases II. Proc. Natl. Acad. Sci. USA **69**:3119–3123.
14. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
15. **Cho, K. W. Y., K. Khalili, R. Zandomeni, and R. Weinmann.** 1985. The gene encoding the large subunit of human RNA polymerase II. J. Biol. Chem. **260**:15204–15210.
16. **Cochet-Meilhac, M., and P. Chambon.** 1974. Animal DNA-dependent RNA polymerases. II. Mechanism of the inhibition of RNA polymerase B by amatoxins. Biochem. Biophys. Acta **353**:160–184.
17. **Corden, J. L., D. L. Cadena, J. M. Ahearn, Jr., and M. E. Dahmus.** 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proc. Natl. Acad. Sci. USA **82**:7934–7938.
18. **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.
19. **Coulson, A., R. Waterston, J. Kiff, J. Sulston, and Y. Kohara.** 1988. Genome linking with yeast artificial chromosomes. Nature (London) **335**:184–186.
20. **Cox, G. N., J. M. Kramer, and D. Hirsh.** 1984. Number and organization of collagen genes in *Caenorhabditis elegans*. Mol. Cell. Biol. **4**:2389–2395.
21. **Deininger, P. L.** 1983. Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. Anal. Biochem. **129**:216–223.
22. **Emmons, S. W., M. R. Klass, and D. Hirsh.** 1979. Analysis of the constancy of DNA sequences during development and evolution of the nematode *C. elegans*. Proc. Natl. Acad. Sci. USA **76**:1333–1337.
23. **Ferguson, E. L., and H. R. Horvitz.** 1985. Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. Genetics **110**:17–72.
24. **Files, J. G., S. Carr, and D. Hirsh.** Actin gene family of *Caenorhabditis elegans*. J. Mol. Biol. **164**:355–375.
25. **Greenleaf, A. L.** 1983. Amanitin-resistant RNA polymerase II

mutations are in the enzyme's largest subunit. J. Biol. Chem. **258**:13403–13406.
26. **Greenleaf, A. L., L. M. Borsett, P. F. Jiamachello, and D. E. Coulter.** 1979. α-Amanitin-resistant *D. melanogaster* with an altered RNA polymerase II. Cell **18**:613–622.
27. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**:3961–3965.
28. **Horvitz, H. R., S. Brenner, J. Hodgkin, and R. K. Herman.** 1979. A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Mol. Gen. Genet. **175**:129–133.
29. **Ingles, C. J., J. Biggs, J. K.-C. Wong, J. R. Weeks, and A. L. Greenleaf.** 1983. Identification of a structural gene for a RNA polymerase II polypeptide in *Drosophila melanogaster* and mammalian species. Proc. Natl. Acad. Sci. USA **80**:3396–3400.
30. **Ingles, C. J., A. Guialis, J. Lam, and L. Siminovitch.** 1976. α-Amanitin resistance of RNA polymerase II in mutant Chinese hamster ovary cell lines. J. Biol. Chem. **251**:2729–2734.
31. **Ingles, C. J., H. J. Himmelfarb, M. Shales, A. L. Greenleaf, and J. D. Friesen.** 1984. Identification, molecular cloning, and mutagenesis of *Saccharomyces cerevisiae* RNA polymerase genes. Proc. Natl. Acad. Sci. USA **81**:2157–2161.
32. **Jokerst, R. S., J. R. Weeks, W. A. Zehring, and A. L. Greenleaf.** 1989. Analysis of the gene encoding the largest subunit of RNA polymerase II in *Drosophila*. Mol. Gen. Genet. **215**:266–275.
33. **Kenyon, C.** 1988. The nematode, *Caenorhabditis elegans*. Science **240**:1448–1453.
34. **Kim, W.-Y., and M. E. Dahmus.** 1989. The major late promoter of adenovirus-2 is accurately transcribed by RNA polymerases IIO, IIA, and IIB. J. Biol. Chem. **264**:3169–3176.
35. **Kramer, J. M., G. N. Cox, and D. Hirsh.** 1982. Comparisons of the complete sequence of two collagen genes from *Caenorhabditis elegans*. Cell **30**:599–606.
36. **Kramer, J. M., G. N. Cox, and D. Hirsh.** 1985. Expression of the *Caenorhabditis elegans* collagen genes *col-1* and *col-2* is developmentally regulated. J. Biol. Chem. **260**:1945–1951.
37. **Lampe, M., M. Igo, W. Schafer, C. Binnie, R. Losick, C. Ray, and C. Moran.** 1987. Studies on the structural gene for σ^{37}, a secondary RNA polymerase sigma factor in *Bacillus subtilis*, p. 177–183. *In* W. S. Reznikoff, C. A. Gross, R. R. Burgess, M. T. Record, Jr., J. E. Dahlberg, and M. P. Wickens (ed.), RNA polymerase and the regulation of transcription. Elsevier Science Publishing, Inc., New York.
38. **Maniatis, T., E. F. Fritch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **Nonet, M., D. Sweetser, and R. A. Young.** 1987. Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. Cell **50**:909–915.
40. **Politz, J. C., and R. S. Edgar.** 1984. Overlapping stage-specific sets of numerous small collagenous polypeptides are translated *in vitro* from *Caenorhabditis elegans* RNA. Cell **37**:853–860.
41. **Riva, M., S. Memet, J.-Y. Micouin, J. Huet, I. Treich, J. Dassa, R. Young, J.-M. Buhler, A. Sentenac, and P. Fromageot.** 1986. Isolation of structural genes for yeast RNA polymerases by immunological screening. Proc. Natl. Acad. Sci. USA **83**:1554–1558.
42. **Rogalski, T. M., A. M. E. Bullerjahn, and D. L. Riddle.** 1988. Lethal and amanitin-resistance mutations in the *Caenorhabditis elegans ama-1* and *ama-2* genes. Genetics **120**:409–422.
43. **Rogalski, T. M., and D. L. Riddle.** 1988. A *Caenorhabditis elegans* RNA polymerase II gene, *ama-1*, and nearby essential genes. Genetics **118**:61–74.
44. **Sanford, T., M. Golomb, and D. L. Riddle.** 1983. RNA polymerase II from wild type and α-amanitin-resistant strains of *Caenorhabditis elegans*. J. Biol. Chem. **258**:12804–12809.
45. **Sanford, T., J. P. Prenger, and M. Golomb.** 1985. Purification and immunological analysis of RNA polymerase II from *Caenorhabditis elegans*. J. Biol. Chem. **260**:8064–8069.
46. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
47. **Searles, L. L., R. S. Jokerst, P. M. Bingham, R. A. Voelker, and A. L. Greenleaf.** 1982. Molecular cloning of sequences from a *Drosophila* RNA polymerase locus by P element transposon tagging. Cell **31**:585–592.
48. **Sentenac, A.** 1985. Eukaryotic RNA polymerases. Crit. Rev. Biochem. **18**:31–91.
49. **Sommers, D. G., M. L. Pearson, and C. J. Ingles.** 1975. Isolation and characterization of an α-amanitin-resistant rat myoblast mutant cell line possessing an α-amanitin-resistant RNA polymerase II. J. Biol. Chem. **250**:4825–4831.
50. **Sulston, J., and S. Brenner.** 1974. The DNA of *Caenorhabditis elegans*. Genetics **77**:95–104.
51. **Sweetser, D., M. Nonet, and R. A. Young.** 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc. Natl. Acad. Sci. USA **84**:1192–1196.
52. **Thomas, P.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
53. **von Mende, N., D. M. Bird, P. S. Albert, and D. L. Riddle.** 1988. *dpy-13*: a nematode collagen gene that affects body shape. Cell **55**:567–576.
54. **Wahl, G. M., M. Stern, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulphate. Proc. Natl. Acad. Sci. USA **76**:3683–3687.
55. **Wood, W. B. (ed.).** 1988. The Nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
56. **Yura, T., and A. Ishihama.** 1979. Genetics of bacterial RNA polymerase. Annu. Rev. Genet. **13**:59–97.
57. **Zehring, W. A., J. M. Lee, J. R. Weeks, R. S. Jokerst, and A. L. Greenleaf.** 1988. The C-terminal repeat domain of RNA polymerase II largest subunit is essential *in vivo* but is not required for accurate transcription initiation *in vitro*. Proc. Natl. Acad. Sci. U.S.A. **85**:3698–3702.