RNA (1995), *1*:293–303. Cambridge University Press. Printed in the USA.

Caenorhabditis elegans embryos contain only one major species of Ro RNP

DEBRA J. VAN HORN,[1]* DAN EISENBERG,[2]* CHARLES A. O'BRIEN,[1] and SANDRA L. WOLIN[1]

[1] Department of Cell Biology and [2] Department of Molecular Biophysics and Biochemistry, and Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA

ABSTRACT

In virtually all vertebrate cells, Ro RNPs consist of the 60-kDa Ro autoantigen bound to one of several small cytoplasmic RNA molecules known as Y RNAs. Because the 60-kDa Ro autoantigen is also found complexed with defective precursors of 5S rRNA in *Xenopus* oocytes, we have proposed that this protein functions in a quality control, or discard pathway, for 5S RNA biosynthesis (O'Brien CA, Wolin SL, 1994, *Genes & Dev 8*:2891–2903). The role of the Y RNAs in this pathway is unknown. To begin a genetic analysis of Ro RNP function, we have characterized these particles in the nematode *Caenorhabditis elegans*. The *C. elegans* Ro protein is 12 kDa larger than the vertebrate protein; the larger size is due in part to an N-terminal extension and to two insertions in the RNA recognition motif. In contrast to all previously described vertebrate species, the Ro protein appears bound to a single Y RNA in *C. elegans*. Similar to vertebrate Y RNAs, the *C. elegans* Y RNA can be folded to form a pyrimidine-rich internal loop and a long stem in which the 5′ and 3′ ends are base paired. Within the stem is a conserved bulged helix that is proposed to be the binding site of the Ro protein. Interestingly, although the human protein can bind the nematode Y RNA, the *C. elegans* protein does not bind human Y RNAs. This is the first description of Ro RNPs in an invertebrate species.

Keywords: nematode; RNA polymerase III; RNA recognition motif; Ro autoantigen; Y RNA

INTRODUCTION

Most vertebrate cells contain a class of small RNA–protein complexes known as Ro ribonucleoproteins (RNPs). These RNPs were initially discovered because they are recognized by anti-Ro antibodies from patients suffering from two rheumatic disorders: systemic lupus erythematosus and Sjogren's syndrome (Mattioli & Reichlin, 1974; Alspaugh & Tan, 1975; Lerner et al., 1981b). Ro RNPs consist of a 60-kDa protein complexed with one of several small cytoplasmic RNA molecules known as Y RNAs (Wolin & Steitz, 1984). cDNAs encoding the 60-kDa protein have been cloned from humans and *Xenopus* (Deutscher et al., 1988; O'Brien et al., 1993). The protein is a member of the family of RNA-binding proteins that contain an ~80-amino acid motif known as an RNA-recognition motif (RRM) or RNA-binding domain (Birney et al., 1993). Although a 52-kDa protein has been proposed to be an additional component of Ro RNPs in human cells (Ben-Chetrit et al., 1988), this is controversial (Kelekar et al., 1994).

The number of distinct Y RNAs associated with the 60-kDa Ro protein varies between different vertebrate species. Although both humans and *Xenopus* cells contain four distinct Y RNA species, mouse cells contain only two Y RNAs (Hendrick et al., 1981; O'Brien et al., 1993). Y RNAs have been sequenced from human, *Xenopus*, and iguana cells; these RNAs are transcribed by RNA polymerase III and range in size from 69 to 112 nt (Wolin & Steitz, 1983; O'Brien et al., 1993; Farris et al., 1995). All the sequenced Y RNAs can be drawn as structures containing a pyrimidine-rich internal loop and a long stem formed by base pairing the 5′ and 3′ ends. Each Ro RNP is present in vertebrate cells at approximately 1% the level of ribosomes.

In addition to binding Y RNAs, the 60-kDa Ro protein was recently shown to be complexed with certain variant 5S rRNAs in *Xenopus* oocytes (O'Brien & Wolin, 1994). These 5S rRNAs contained extra nucleotides at the 3′ end, and were apparently generated by read-through of the first transcription termination signal. All

Reprint requests to: Sandra L. Wolin, Department of Cell Biology and Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, Connecticut 06510, USA; e-mail: sandra_wolin@quickmail.yale.edu.

* The first two authors contributed equally to the work.

the complexed 5S rRNAs also contained internal point mutations. Because these mutant RNAs were inefficiently processed to mature 5S rRNA and were eventually degraded, the 60-kDa protein has been proposed to participate in a quality control pathway for 5S rRNA production (O'Brien & Wolin, 1994). It is unclear what role the Y RNAs might play in such a pathway.

The identification of Ro RNPs in genetically tractable organisms would greatly facilitate dissection of Ro RNP function. This would allow us to assess the consequences of genetically depleting these particles from cells or organisms. It would also be possible to identify factors that interact with Ro RNPs by obtaining extragenic suppressor mutations. To begin a genetic analysis of Ro RNP function, we have characterized Ro RNPs in the nematode *Caenorhabditis elegans*. A combination of in vitro and in vivo analyses of splicing in nematodes has yielded a wealth of information as to the biology of pre-mRNA splicing (reviewed by Nilsen, 1993). In addition, the extensive analyses of development and cell lineage that have been performed in *C. elegans* make it especially attractive for in vivo analyses of cellular processes (Sulston et al., 1983). We report here that Ro RNPs, which have previously been identified only in vertebrates, are also present in *C. elegans*. In contrast to all previously characterized vertebrate species, we have detected only a single Y RNA associated with the Ro protein in this nematode. A bulged helix that has been implicated as the binding site of the 60-kDa Ro protein is the most conserved feature of the *C. elegans* Y RNA. Although the human protein can bind both the human and worm Y RNAs, the *C. elegans* protein is unable to bind human Y RNAs, indicating that at least in worms, other structural features are important for Ro protein binding.

RESULTS

A homologue of the 60-kDa Ro protein in *C. elegans*

As part of the *C. elegans* genome project, a collection of cDNAs was generated, and a single sequencing run was performed from the 5′ end of each clone (Waterston et al., 1992). One of these sequences (cDNA clone cm11g4, 371 nt) could encode a protein containing an N-terminus similar to the human and *Xenopus* 60-kDa Ro proteins. We completed the sequence of the cDNA clone and found that it could encode an open reading frame of 643 amino acids with a predicted molecular mass of 72.8 kDa (Fig. 1). Because the open reading frame did not contain a stop codon upstream of the first AUG, we used an internal primer to amplify the 5′ ends of additional cDNAs from a cDNA library prepared from adult worms (Barstead & Waterston, 1989). In addition to containing a stop codon, the first 9 nt of the amplified sequence were identical to the 3′ end of the SL1 spliced leader sequence (Krause & Hirsh, 1987).

The sequence of the worm protein is compared with that of the previously sequenced human and *Xenopus* Ro proteins in Figure 2. The three proteins are 36% identical overall, with the identity extending throughout the length of the protein. Much of the larger size of the worm protein is due to an extension of 47 amino acids at the N-terminus. Interestingly, the worm sequence also contains two insertions, one of 19 amino acids and a second of 6 amino acids, in the RNA recognition motif. We note that a potential zinc finger motif noted in the human sequence (Deutscher et al., 1988) is not conserved in either the *Xenopus* or worm proteins (amino acids 411–429 in Fig. 2).

The cm11g4 cDNA has been mapped to chromosome V, between *her-1* and *act-1* (ACEDB, the *C. elegans* genome database; R. Waterston, pers. comm.).

A single Y RNA is bound to the *C. elegans* Ro protein in embryo extracts

To characterize the small RNAs bound by the worm protein, we prepared rabbit antibodies against a fusion protein containing the potential Ro protein homologue linked to polyhistidine. When these antibodies were used in Western blotting experiments against extracts of adult worms, a single polypeptide of ~69 kDa was detected (Fig. 3A, lane 1). Although this polypeptide was slightly smaller than the predicted molecular size, we obtained a protein of similar size when we subjected the cDNA clone to in vitro transcription and translation (data not shown).

To determine whether the worm protein, like the vertebrate 60-kDa Ro proteins, was bound to Y RNAs in vivo, we performed immunoprecipitations from extracts of *C. elegans* embryos and examined the RNAs contained within the immunoprecipitates. The RNAs were visualized by labeling with ^{32}P-pCp. Using the rabbit antibody against the *C. elegans* Ro protein, one major small RNA was present in the immunoprecipitates (Fig. 3B, lane 3). This RNA was absent when the immunoprecipitation was performed using preimmune sera (lane 4). As a control, we also performed immunoprecipitations with a monoclonal antibody, Y-12, which recognizes the Sm proteins that are components of the spliceosomal U snRNPs (Lerner et al., 1981a). Although this antibody has been reported not to cross-react with *C. elegans* U RNPs (Van Doren & Hirsh, 1988), it efficiently immunoprecipitated these RNPs in our experiments (lane 2).

We also attempted to immunoprecipitate Ro RNPs from sonicates of adult worms. These extracts contained a great deal of ribonuclease, and there were many bands that appeared in all the immunoprecipitates, including those with preimmune sera (data not shown). However, a band of approximately 100 nt ap-

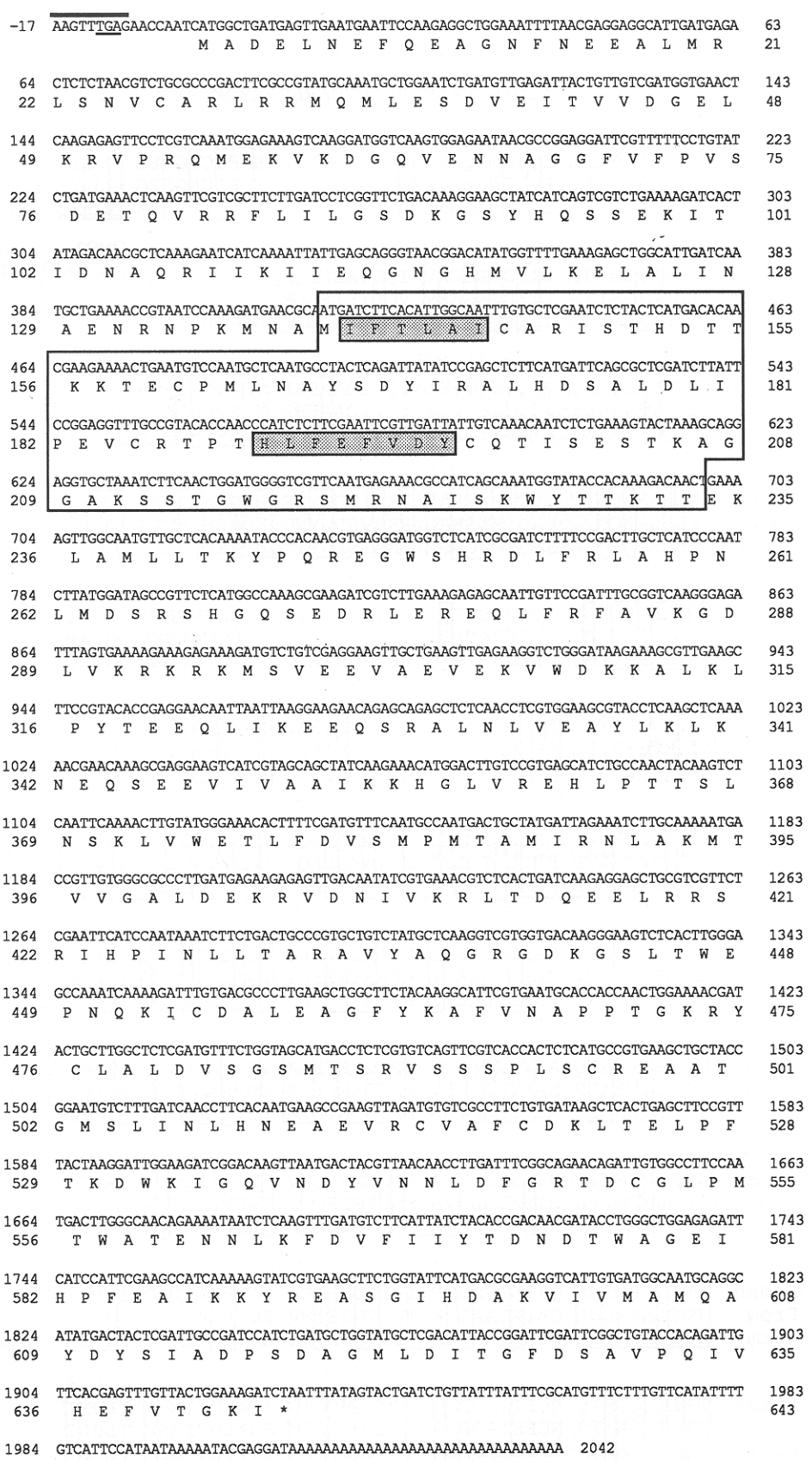

FIGURE 1. Nucleotide and predicted amino acid sequence of the *C. elegans* 60-kDa Ro protein. An in-frame stop codon upstream of the first AUG is underlined. This stop codon is contained within a 9-nt sequence that is identical to 3′ end of the SL1 spliced leader sequence (Krause & Hirsh, 1987; indicated by a line above the sequence). The RRM is boxed, and two conserved submotifs within the RRM—RNP1 and RNP2—are shaded. The sequence of the cDNA encoding the nematode Ro protein has been assigned GenBank accession number L41729.

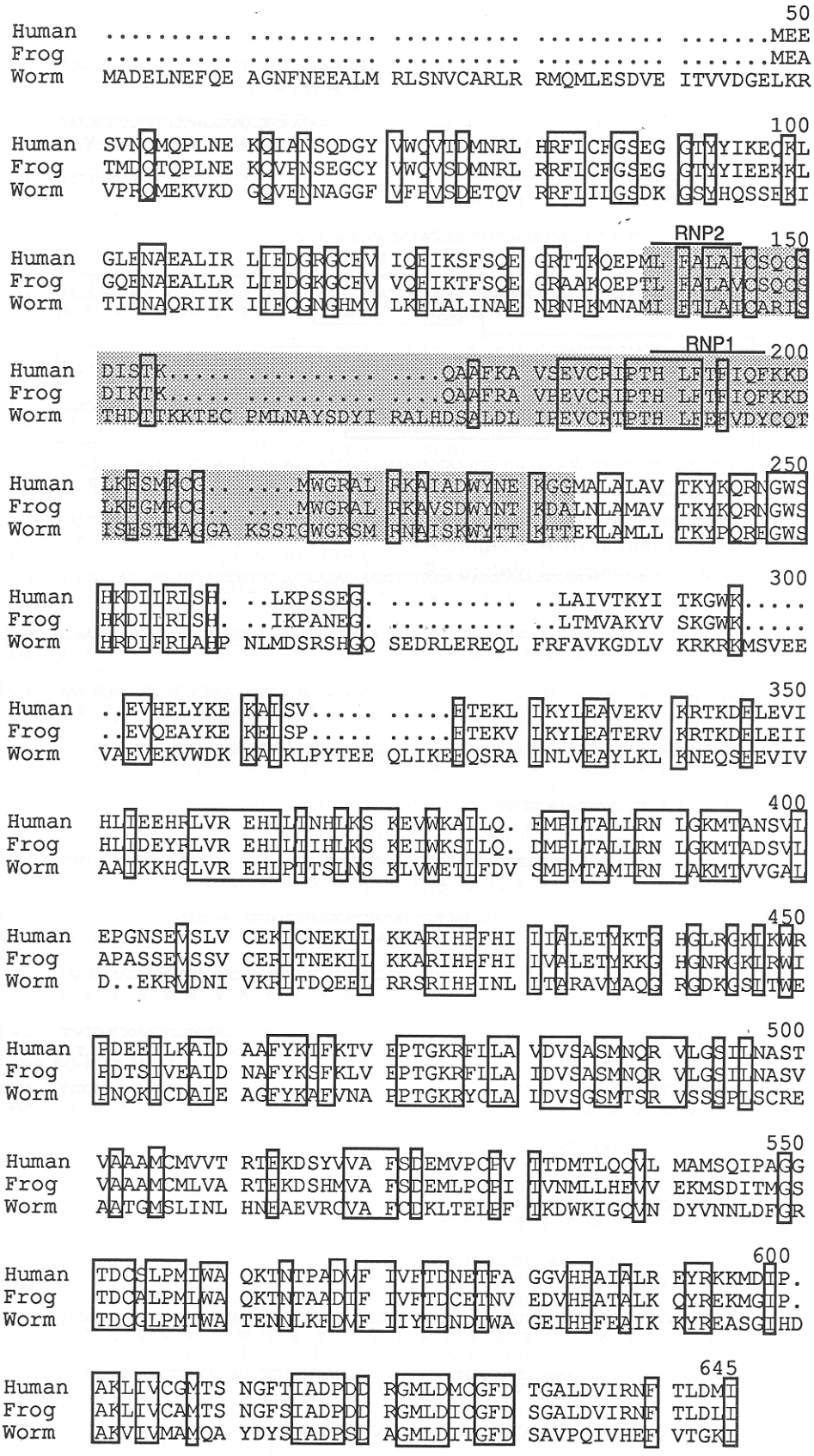

FIGURE 2. Comparison of the human, *Xenopus*, and *C. elegans* 60-kDa Ro proteins. The three Ro proteins were aligned using the GCG program PILEUP. Amino acids that are identical between the three proteins are boxed. Gaps inserted to maximize homology are indicated by dots. RNA recognition motif is shaded, and RNP1 and RNP2 submotifs are indicated by lines above the sequences.

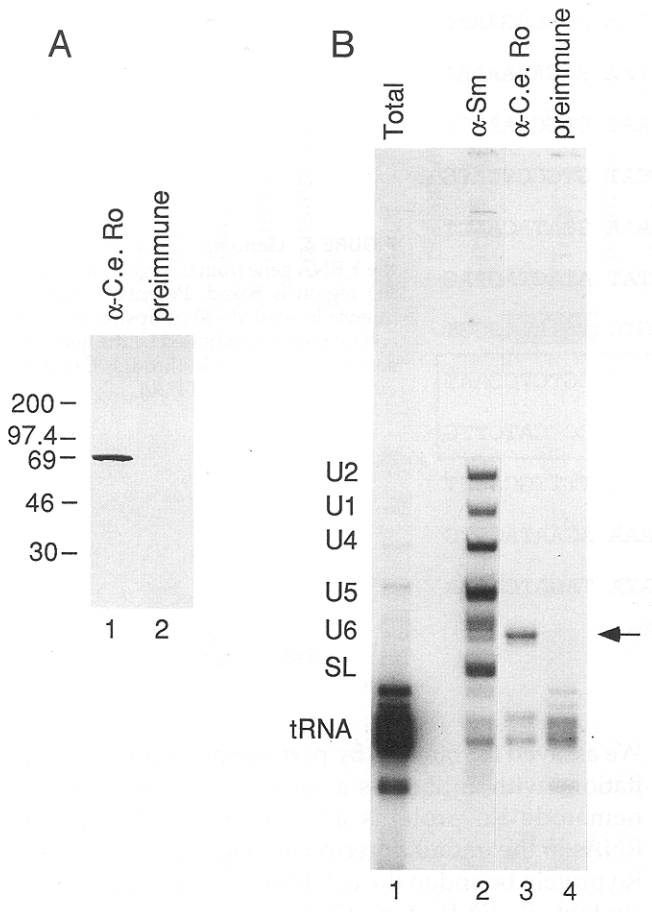

FIGURE 3. Antibodies against the *C. elegans* 60-kDa Ro protein immunoprecipitate a single Y RNA from worm extracts. **A:** Extracts of adult worms were subjected to immunoblotting using either the antibody to the *C. elegans* 60-kDa Ro protein (lane 1) or preimmune serum (lane 2). **B:** Embryo extracts were prepared as described in the Materials and methods and subjected to immunoprecipitation with either monoclonal anti-Sm antibodies (lane 2), anti-*C. elegans* Ro antibodies (lane 3), or preimmune sera (lane 4). RNAs contained within immunoprecipitates were labeled with [^{32}P]pCp and fractionated on a denaturing polyacrylamide gel. The total RNA profile (lane 1) was prepared by phenol extraction of a small fraction of the lysate. Because the majority of the ribosomes were pelleted during extract preparation (see Materials and methods), 5S and 5.8S RNA are underrepresented in the total RNA profile (lane 1). Identities of the RNAs in the anti-Sm immunoprecipitate were inferred by comparing their mobilities with the known sizes of the *C. elegans* snRNAs (Thomas et al., 1988, 1990).

peared reproducibly enriched in the immunoprecipitates. Partial sequence analysis of this band revealed that it was identical to the major RNA present in the embryo immunoprecipitates (data not shown).

The sequence of the prominent RNA in the anti-Ro immunoprecipitate was determined by a combination of enzymatic and primer extension sequencing. Ambiguities were resolved by using an oligonucleotide complementary to 29 nt at the 3′ end to obtain a genomic clone. The sequence of the genomic DNA was completely consistent with the RNA sequence and is shown in Figure 4. As was previously found for the genes encoding the *Xenopus* Y RNAs (O'Brien et al., 1993), a TATA-like sequence is located 25–30 nt upstream of the coding region.

The sequence of the nematode RNA is compared with the four human Y RNA sequences in Figure 5. Like the previously sequenced vertebrate Y RNAs, the nematode RNA can be drawn as a structure in which the 5′ and 3′ ends are base paired to form a long stem. Within this stem is a structural motif that has been noted in all Y RNAs sequenced to date, a single bulged cytidine with a conserved helix (boxed in Fig. 5). In vertebrates, this helix consists of seven nearly identical base pairs that differ only in that the first U-A base pair is replaced by a U-G in Y4 RNAs. In the *C. elegans* Y RNA, only five of the base pairs are completely preserved, because the first base pair is not conserved and the next G-U base pair is replaced by a G-C. The stem also differs in the nematode RNA, in that the base pairing is interrupted by a dinucleotide bulge near the base of the stem. Similar to the vertebrate Y RNAs, the nematode RNA structure contains a large pyrimidine rich internal loop. The loop shares an 8-nt sequence, YUU CUUUR with the human, *Xenopus*, and iguana Y3 RNAs (indicated by line in Fig. 5). The nematode loop also contains a dinucleotide repeat, the sequence (UG) repeated eight times.

The single small RNA bound by the nematode Ro protein homologue is clearly a *C. elegans* Y RNA because it contains several structural features in common with the previously characterized vertebrate Y RNAs. Although it contains a sequence in the pyrimidine-rich internal loop common to vertebrate Y3 RNAs, the *C. elegans* Y RNA is not otherwise more related to Y3 RNA than to several of the other vertebrate Y RNAs. Because we do not yet know whether this single Y RNA is a functional homologue of Y3 RNA, we refer to the RNA simply as the *C. elegans* Y RNA (ceY RNA).

To determine the location of the gene encoding the ceY RNA, we synthesized the RNA in vitro with T7 polymerase and hybridized it to a filter of yeast artificial chromosome (YAC) DNAs containing genomic sequences that span the nematode genome (Coulson et al., 1988). In addition, the λ clone containing the ceY RNA gene was sent to Dr. A. Coulson to be placed on the physical map of the *C. elegans* genome using DNA fingerprinting techniques (Coulson et al., 1986). Both mapping experiments localized the Y RNA gene to chromosome IV, between the *lin-45* and *col-4* genes. The λ clone has been named NH#LYR (ACEDB).

The nematode Ro protein assembles with the *C. elegans* Y RNA to form RNPs, but does not bind human hY3 RNA

Because both the protein and RNA components of Ro RNPs appear to have significantly diverged between nematodes and humans, we determined if the differ-

```
  1  CTGACATAAA AATTCAAAAA ATCACACCTA TAAAATTGCA ACACGGAATC
 51  GGAAAACCGA AATTGATAAT GGAACTTTTG GAGAAAATAA AAGAAAAAAT
101  ATTTAAATTC ACAAATGGAA AATGAACCAA CCAGTTTAAT GAGCAAAACC
151  ACACAACATC CAAAAAGAAT TGAAATATCT CCTCTTGCAT GTGCCTTATG
201  ATGAGCATCT AATCTTTCAT TTGCAATTGT TATCGAAAAA GCATACAACT
251  GAAAATGTTT TATGTTGGAA GAAAATATCT CACACAATAT ATAGTAGAAG
301  GCCAAACTCG GAAAATGTCG GGCACTTTCC TATGCAAGTG TATATAGGGT
351  TAGTTTATGT TTAAACATTT GGGCTCGGTC CGAGTTTCAT GGTCTCCAAT
401  GTGTGTGTGT GTGTGTTTTC TTTAGGAACC TCGGTTCCAA CCTCATCTTG
451  ACCTTGAAAC TACTTTGACC GCTCCTTTTG GATTTCCGAG TTTTGCACTT
501  TTTAACTTTT AAAATGTCAA TGTACTTACG GCTCCGAAAA AGAATACTTG
551  ACAAATGTTT GCAGCCAAAT TGATATAGTC ATTTGTCATA TAGATCATCA
601  ACTAGCTCAT CG
```

FIGURE 4. Genomic DNA sequence of the Y RNA gene from *C. elegans*. RNA coding region is boxed. Potential TATA sequence located 25–30 nt upstream of the coding region is indicated by the line. This sequence is available through GenBank (accession number L14730).

ent Ro proteins would be able to form chimeric Ro RNPs with Y RNAs from the other species. We translated synthetic mRNAs encoding these proteins in a rabbit reticulocyte lysate and examined the ability of the translated protein to assemble with ^{32}P-labeled Y RNAs. As a negative control, we included the nucleolar U3 RNA. (Both Y RNAs, as well as the U3 RNA, were synthesized in vitro using T7 RNA polymerase.) We assayed for binding by performing immunoprecipitations with antibodies against either the human or nematode Ro proteins and examining the labeled RNAs in the immunoprecipitate (Fig. 6). The *C. elegans* Ro protein bound to the ceY RNA to form immunoprecipitable Ro RNPs (lane 12), but did not bind to either the human Y3 RNA (lane 9) or to hY4 RNA (data not shown). However, the human protein assembled with

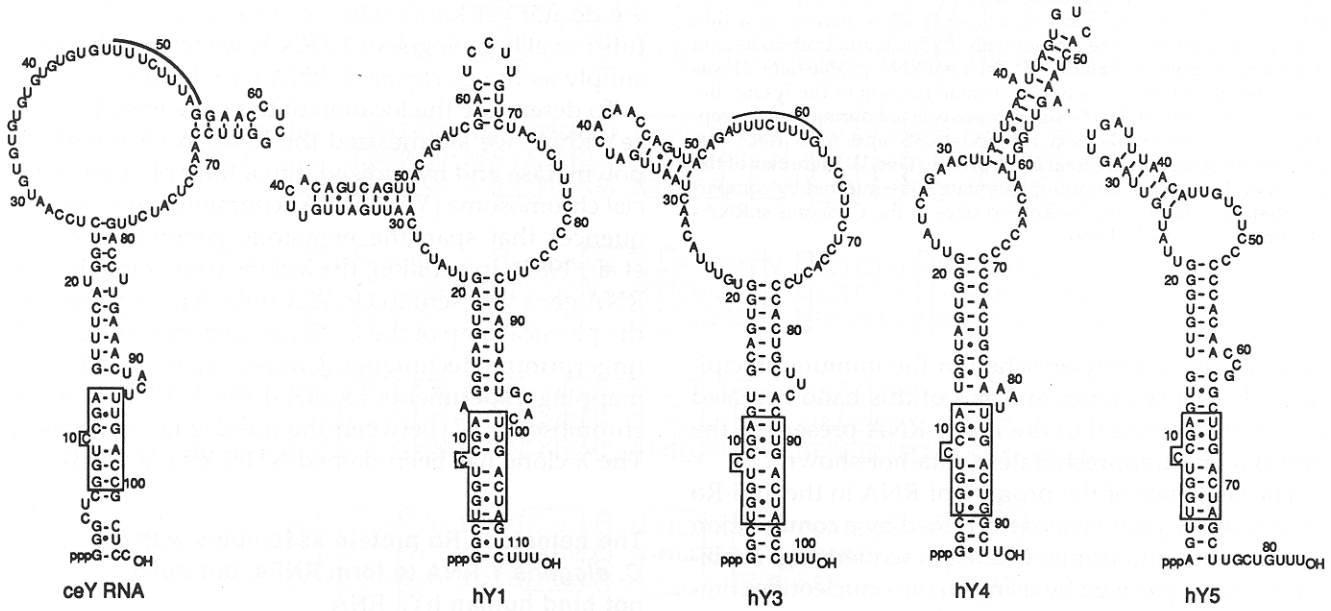

FIGURE 5. Potential secondary structures of *C. elegans* and human Y RNAs. Structures of the four human Y RNAs were proposed by O'Brien et al. (1993) and were drawn to maximize structural similarities between the human and *Xenopus* Y RNAs. A conserved helix that was proposed to be the binding site for the 60-kDa Ro protein is boxed (Wolin & Steitz, 1984). A sequence in the loop of ceY RNA that is also present in the human, *Xenopus*, and iguana Y3 RNAs is indicated by the line.

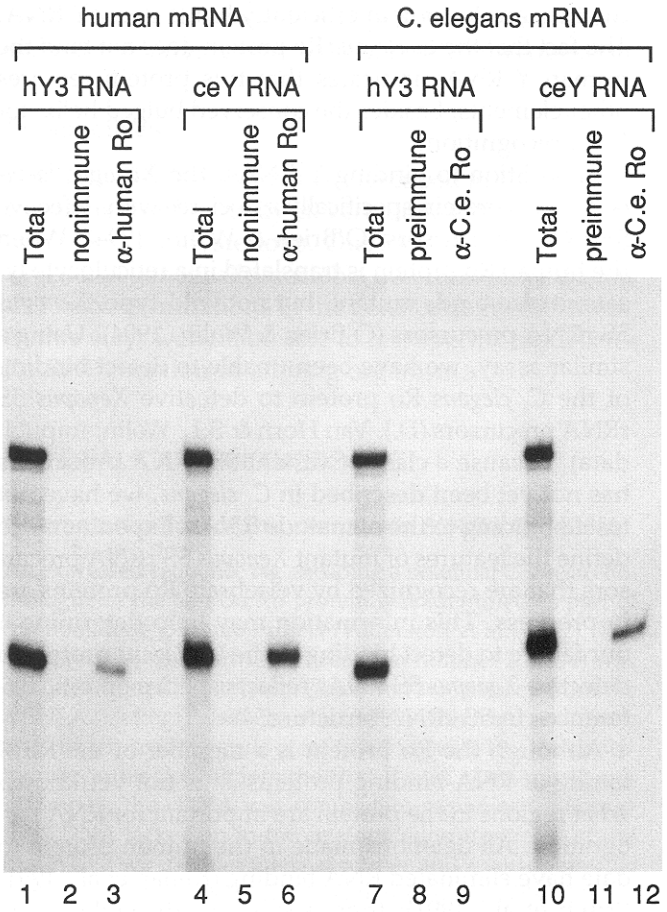

FIGURE 6. Assembly of Ro RNPs. Synthetic mRNAs encoding the human (lanes 1–6) or *C. elegans* (lanes 7–12) Ro proteins were translated in a rabbit reticulocyte lysate. Following translation, the lysate was incubated with a mixture of ^{32}P-labeled U3 RNA and either the human Y3 (lanes 1–3 and 7–9) or the *C. elegans* Y RNA (lanes 4–6 and 10–12). The extract was then divided into equal aliquots and either extracted with phenol (lanes 1, 4, 7, 10) or subjected to immunoprecipitation with either a nonimmune human serum (lanes 2, 5), a patient anti-Ro serum (lanes 3, 6), rabbit preimmune serum (lanes 8, 11), or rabbit anti-*C. elegans* Ro antibodies (lanes 9, 12). RNAs present in the immunoprecipitates were extracted with phenol and fractionated in a 5% polyacrylamide–8 M urea gel.

both hY3 RNA and ceY RNA to form immunoprecipitable RNPs (lanes 3, 6). (This represents binding of the human protein, rather than free Ro protein present in the reticulocyte lysate, to the added RNA, as we previously found that assembly of human Ro RNPs in this assay requires mRNA translation [O'Brien & Wolin, 1994].) Somewhat surprisingly, the human protein reproducibly bound more of the added *C. elegans* Y RNA than the human Y RNA (Fig. 6, compare lanes 3 and 6).

DISCUSSION

In vertebrate cells, Ro RNPs consist of the 60-kDa Ro protein bound to one of several small RNAs known as Y RNAs. Although Ro RNPs are abundant and ubiquitous components of vertebrate cells, their function remains unknown. The 60-kDa Ro protein is also found complexed with a class of variant 5S rRNA precursors in *Xenopus* oocytes. Because these variant 5S rRNA precursors are processed inefficiently to 5S rRNA and most are eventually degraded, the Ro protein has been proposed to function in a quality control, or discard pathway, for 5S rRNA biosynthesis (O'Brien & Wolin, 1994). The role that Y RNAs play in this pathway is not known.

As a first step in a genetic analysis of Ro RNP function, we have characterized Ro RNPs in the nematode *C. elegans*. These experiments have revealed that Ro RNPs are not restricted to vertebrate cells, but also occur in invertebrates. In contrast to all previously characterized vertebrate species, we have detected only one species of Y RNA in *C. elegans*. The most conserved feature of the ceY RNA is a bulged helix that has been proposed to be the binding site for the Ro protein on the Y RNAs. Although the human protein can bind the ceY RNA, the *C. elegans* Ro protein does not bind human Y RNAs, indicating that structural features required for binding have diverged between humans and worms.

A single Y RNA species in *C. elegans*

In vertebrates, the number of distinct Y RNAs associated with the Ro 60-kDa protein varies from two to four depending on the species examined. The fact that *C. elegans* apparently contain only a single major Y RNA suggests that the multiple Y RNAs found in vertebrate species arose by duplication and divergence of a single ancestral gene. This idea is supported by the fact that the genes encoding all four human Y RNAs appear to be tightly linked on human chromosome 7 (Wolin & Steitz, 1983; Maraia et al., 1994). We do not yet know

if the additional Y RNAs in vertebrates carry out functions that are unnecessary in *C. elegans*. An alternative possibility is that the ceY RNA performs all roles carried out in vertebrates by multiple Y RNA species.

Although our most successful immunoprecipitations were performed in embryonic extracts, we consider it unlikely that adult worms contain additional Y RNAs for several reasons. First, a variety of differentiated cell types are present in *C. elegans* embryos (Sulston et al., 1983). Second, although human erythrocytes contain only two of the four human Y RNAs (O'Brien & Harley, 1990), the majority of vertebrate tissues, including *Xenopus* oocytes, contain the expected number of Y RNAs (Pruijn et al., 1993; O'Brien & Wolin, 1994; S. Wolin, unpubl. obs.). However, due to difficulties in performing immunoprecipitations from adult worms, we cannot rule out the possibility that *C. elegans* contains additional Y RNA species that have escaped our detection.

Our identification of the ceY RNA has allowed us to greatly extend the existing phylogenetic comparisons of Y RNA structure. All the vertebrate Y RNAs sequenced to date share a long stem, in which the 5′ and 3′ ends are base paired, and a pyrimidine-rich internal loop. Experimental support that both the stem and loop exist has been provided by enzymatic and chemical modification of naked hY1 and hY5 RNAs (van Gelder et al., 1994) as well as by oligonucleotide-directed RNase H digestion of native human RNPs (Matera et al., 1995). The *C. elegans* Y RNA can be folded to form a similar structure. An unusual feature of the large internal loop in the worm RNA is that it contains the dinucleotide GU repeated eight times. The primary sequence of this loop is most related to that of the vertebrate Y3 RNAs, in that all Y3 RNAs characterized to date contain the sequence YUUCUUUR in this loop. Because Y3 RNA is the most conserved Y RNA in vertebrates (O'Brien et al., 1993; Farris et al., 1995), it is possible that the *C. elegans* Y RNA is a Y3 homologue. However, the definitive classification of this RNA as a Y3 RNA will require the identification of the function(s) of the multiple Y RNAs in vertebrate species.

Binding of the Ro 60-kDa protein to RNA

Although the features of Y RNA structure that determine binding by the 60-kDa Ro protein have not been precisely defined, the conserved bulged helix is within the region protected by bound human Ro protein from nuclease digestion (Wolin & Steitz, 1984). Mutagenesis experiments have confirmed that the bulged cytidine within the helix is critical for Ro protein recognition (Pruijn et al., 1991). The *C. elegans* Y RNA possesses a very similar helix, which differs only slightly from the mammalian motif (boxed in Fig. 5). As would be expected if this structure were sufficient for recognition, the human Ro protein efficiently binds the ceY RNA. The fact that the *C. elegans* Ro protein does not bind the human Y RNAs indicates that this protein requires other elements, besides the conserved bulged helix, for RNA recognition.

In addition to binding Y RNAs, the *Xenopus laevis* 60-kDa Ro protein specifically associates with defective 5S rRNA precursors (O'Brien & Wolin, 1994). When the human Ro protein is translated in a reticulocyte lysate, it also binds mutant, but not wild-type, *Xenopus* 5S rRNA precursors (O'Brien & Wolin, 1994). Using a similar assay, we have been unable to detect binding of the *C. elegans* Ro protein to defective *Xenopus* 5S rRNA precursors (D.J. Van Horn & S.L. Wolin, unpubl. data). Because a class of variant 5S rRNA precursors has not yet been described in *C. elegans*, we have not tested binding to the nematode RNAs. Experiments to define the features of mutant *Xenopus* 5S rRNA precursors that are recognized by vertebrate Ro proteins are in progress. This information may help determine if our failure to detect binding of the *C. elegans* protein to defective *Xenopus* 5S RNAs reflects species-specific differences in 5S rRNA structure.

Although the Ro protein is a member of the RRM family of RNA-binding proteins, it is not yet known what regions in the protein are important for RNA recognition. All deletions made in the human protein to date have eliminated RNA binding (Kenan et al., 1991; Pruijn et al., 1991); thus, sequences throughout the protein are likely to be critical for formation of the RNA binding site. The fact that the protein is conserved throughout its length is consistent with this idea. Despite the overall conservation of protein sequence, a potential zinc finger noted in the human sequence (Deutscher et al., 1988) is not present in either the frog or worm Ro protein. Thus, there is as yet no phylogenetic evidence to indicate that such a structure actually forms.

Although sequences in addition to the RRM are likely to be required for specific RNA binding, it was unexpected to find insertions within this domain in the *C. elegans* protein. X-ray diffraction and NMR studies have revealed that the RRM consists of a sheet of four antiparallel β-strands flanked by two α-helices (Nagai et al., 1990; Hoffman et al., 1991; Gorlach et al., 1992; Oubridge et al., 1994). Alignment of the Ro protein sequence with these structures suggests that the 19-amino acid insertion may enlarge loop 2, which separates the α_1-helix from the β_2-strand. The second insertion may enlarge the loop between α_2 and β_4. Definitive proof that the insertions in the *C. elegans* protein enlarge these two loops will require determination of the protein structure. However, both of these loops lie on the side of the RRM opposite the RNA-binding face. Thus, if this model is correct, the insertions in the RRM may not account for the altered RNA-binding properties of the *C. elegans* Ro protein. Definition of the regions in

the *C. elegans* Ro protein and Y RNA that are required for specific binding may help to resolve this question.

Perspectives

Although Ro RNPs were first described in 1981 and are components of virtually all vertebrate cells, their function has been mysterious. Our identification of these RNPs in a genetically tractable invertebrate will allow us to isolate mutations in the genes that encode them. In this way, we will be able to test specific models for Ro RNP function in a living organism.

MATERIALS AND METHODS

Characterization of the nematode 60-kDa Ro cDNA

An expressed sequence tag encoding a potential *C. elegans* homologue of the 60-kDa Ro protein was identified as part of the nematode genome project (Waterston et al., 1992). The clone (cDNA clone cm11g4, accession number Z14391) was obtained from R. Waterston. The insert was excised from the λ SHLX2 vector (Palazzolo et al., 1990) with *Not* I and *Apa* I, and the protruding ends were filled in with T4 DNA polymerase. After cloning into the *Hinc* II sites of m13mp18 and m13mp19, the cDNA was completely sequenced. Because this cDNA lacked an in-frame stop codon upstream of the first AUG, we used an internal primer and a vector-specific primer to amplify additional 5′ sequence from a cDNA library in λzapII (Barstead & Waterston, 1989).

Antisera

To prepare antibodies against the *C. elegans* 60-kDa Ro protein, the oligonucleotides 5′-CGCGGGATCCCAAATGGAG AAAGTCAAG-3′ and 5′-GGCGGCCTGCAGTTAGATCTT TCCAGTAAC-3′ were used to amplify a DNA fragment encoding amino acids 54–643 of the predicted protein. The amplified fragment was digested with *Bam*H I and *Pst* I and inserted into the corresponding sites of pTrcHis (Invitrogen). Because the recombinant fusion protein was insoluble, it was solubilized in 6 M guanidine-HCl and purified on a ProBond column (Invitrogen) using the denaturing conditions described by the manufacturer. Because the purified protein precipitated out of solution during the dialysis step, we subjected the purified protein to a denaturation and renaturation procedure previously shown to be effective for solubilizing the human 60-kDa Ro protein (M. Saitta & J. Keene, pers. comm.). Briefly, after the protein was eluted from the ProBond resin in 8 M urea, it was dialyzed into 7 M guanidine-HCl, 0.1 M potassium phosphate, pH 7.5, 50 mM DTT, 1 mM PMSF. The protein was diluted 1:10 into renaturation buffer (2.5 M urea, 50 mM Tris, 10 mM NaCl, 5 mM EDTA, 10 mM lysine, 0.5 mM reduced glutathione, 0.3 mM oxidized glutathione) and incubated overnight at 15 °C. The protein was then dialyzed against renaturation buffer containing 1.0 M urea, followed by renaturation buffer without urea. The protein was concentrated by dialysis against dry sucrose, further concentrated in a Speed-Vac (Savant), and used to immunize rabbits.

Immunoprecipitation and immunoblotting

To obtain worm extracts for immunoblotting, wild-type worms (strain N2; a gift of M. Stern, Yale University) were grown on agar plates as described (Sulston & Hodgkin, 1988). Worms were collected from the plates by washing the plates with TBS (40 mM Tris, pH 7.5, 150 mM NaCl). After spinning the worm suspension in a microfuge for 2 min, the worm pellet was washed with water and resuspended in SDS buffer (83 mM Tris base, 30% glycerol, 6.8% SDS, 240 mM DTT). The extract was heated to 100 °C for 5 min, fractionated in an SDS–polyacrylamide gel and transferred to nitrocellulose. Nitrocellulose filters were blocked and probed with antisera as previously described (Yoo & Wolin, 1994).

For immunoprecipitations, embryos were harvested as described by Emmons et al. (1979) and resuspended in NET-2 (20 mM Tris, pH 7.5, 150 mM NaCl, 0.1% NP-40) containing 0.05 mM DTT, 0.5 mM PMSF, 10 mM vanadyl ribonucleoside complexes, and 125 ng/mL each of pepstatin, chymostatin, leupeptin, and antipain. After sonication (three times for 30 s each on setting 4.5), the extract was sedimented at 100,000 × *g* in a Beckman TLA 100.2 rotor for 1 h. Immunoprecipitations were performed as described previously (Wolin & Steitz, 1984). The monoclonal anti-Sm antibody (Y-12) was a gift of M.-D. Shu and J. Steitz.

Sequencing of the *C. elegans* Y RNA

RNA immunoprecipitated from worm embryos was 5′ or 3′ end-labeled with ^{32}P and sequenced with base-specific nucleases as previously described (O'Brien et al., 1993). Additional sequence was obtained by primer extension in the presence of dideoxynucleotides as described by Montzka and Steitz (1988). To resolve ambiguities, we obtained a genomic clone encoding the nematode Y RNA by using the oligonucleotide 5′-AGCGGTCAAAGTGGTTTCAAGGTCAAGAT-3′ to screen a library of *C. elegans* genomic DNA in λ FIX II (Stratagene). Hybridization was performed in 7% SDS, 10% polyethylene glycol 8000, 0.25 M NaCl, 0.13 M sodium phosphate buffer, pH 7.25, at 37 °C. Filters were washed in 5× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.2% SDS at 55 °C for 40 min. From 5×10^5 plaques, four plaques were purified and characterized. They were found to consist of overlapping 15–20-kb inserts, each of which contained a 1.5-kb *Eco*R I fragment that hybridized to the oligonucleotide probe. This fragment was subcloned into the *Eco*R I site of pBluescript II KS− (Stratagene) and sequenced. Although the sequence of this clone contained a single mismatch from the oligonucleotide sequence, the identity of this pyrimidine (the U at position 91 of the RNA coding sequence) had been ambiguous in the direct RNA sequencing data. The 5′ and 3′ terminal nucleotides of the RNA were determined by digestion of end-labeled RNA followed by thin-layer chromatography with appropriate nucleotide standards (Silberklang et al., 1979). In addition, the 5′ nucleotide was confirmed by primer extension.

In vitro translations and RNP reconstitutions

For in vitro translation, the cDNA clone cm11g4 was digested with *Nsi* I and *Apa* I and the *C. elegans* Ro cDNA was inserted into the *Pst* I/*Apa* I sites of pBluescript II KS− (Stratagene).

Following linearization with *Apa* I, transcription with T7 RNA polymerase yielded a synthetic mRNA encoding the *C. elegans* Ro protein.

To synthesize the *C. elegans* Y RNA in vitro, we used PCR to place the *C. elegans* Y RNA coding sequence behind a T7 promoter. The 5′ primer contained an *Eco*R I site, followed by a T7 promoter and 14 nt of the ceY RNA sequence. The 3′ primer contained 16 nt of the ceY RNA sequence, preceded by a *Dra* I site and a *Bam*H I site. Following PCR using the cloned ceY RNA gene as template, the product was digested with *Eco*R I and *Bam*H I and inserted into the *Eco*R I/*Bam*H I site of pSP64 (Promega). After cleavage with *Dra* I, transcription of the clone with T7 RNA polymerase yielded the ceY RNA containing three additional uridine residues at the 3′ end, which probably corresponds to the primary transcript synthesized by RNA polymerase III.

Transcription, translation, and reconstitutions were as previously described (O'Brien & Wolin, 1994), except that immunoprecipitated RNPs were washed with NET-2 rather than HBS-T.

ACKNOWLEDGMENTS

We are grateful to R. Waterston for sending us the cm11g4 cDNA clone, and to M. Stern, L. DeLong, D. DeVore, and the rest of the Stern laboratory for their generosity in providing worms and protocols. We also thank G. Ghosh, J. Keene, M. Saitta, S. Mount, B. Peculis, A. Weiner, and C. Yoo for advice and helpful discussions, and R. Mancebo, S. Mount, and M. Stern for comments on the manuscript. C.A.O. was supported by a postdoctoral fellowship from the Patrick and Catherine Weldon Donaghue Medical Research Foundation. This work was supported by grant GM48410 from the National Institutes of Health.

Received March 27, 1995; returned for revision April 7, 1995; revised manuscript received April 25, 1995

REFERENCES

Alspaugh MA, Tan EM. 1975. Antibodies to cellular antigens in Sjogren's syndrome. *J Clin Invest 55*:1067–1073.

Barstead RJ, Waterston RH. 1989. The basal component of the nematode dense-body is vinculin. *J Biol Chem 264*:10177–10185.

Ben-Chetrit E, Chan EKL, Sullivan KF, Tan EM. 1988. A 52-kD protein is a novel component of the SS-A/Ro antigenic particle. *J Exp Med 167*:1560–1572.

Birney E, Kumar S, Krainer AR. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res 21*:5803–5816.

Coulson A, Sulston J, Brenner S, Karn J. 1986. Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc Natl Acad Sci USA 83*:7821–7825.

Coulson A, Waterston R, Kiff J, Sulston J, Kohara Y. 1988. Genome linking with yeast artificial chromosomes. *Nature 335*:184–186.

Deutscher SL, Harley JB, Keene JD. 1988. Molecular analysis of the 60-kDa human Ro ribonucleoprotein. *Proc Natl Acad Sci USA 85*: 9479–9483.

Emmons SW, Klass MR, Hirsh D. 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.

Farris AD, O'Brien CA, Harley JB. 1995. Y3 is the most conserved small RNA component of Ro ribonucleoprotein complexes in vertebrate species. *Gene 154*:193–198.

Gorlach M, Wittekind M, Beckman RA, Mueller L, Dreyfuss G. 1992. Interaction of the RNA-binding domain of the hnRNP C proteins with RNA. *EMBO J 11*:3289–3295.

Hendrick JP, Wolin SL, Rinke J, Lerner MR, Steitz JA. 1981. Ro small cytoplasmic ribonucleoproteins are a subclass of La ribonucleoproteins: Further characterization of the Ro and La small ribonucleoproteins from uninfected mammalian cells. *Mol Cell Biol 1*: 1138–1149.

Hoffman DW, Query CC, Golden BL, White SW, Keene JD. 1991. RNA-binding domain of the A protein component of the U1 small nuclear ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins. *Proc Natl Acad Sci USA 88*:2495–2499.

Kelekar A, Saitta MR, Keene JD. 1994. Molecular composition of Ro small ribonucleoprotein complexes in human cells. *J Clin Invest 93*:1637–1644.

Kenan DJ, Query CC, Keene JD. 1991. RNA recognition: Towards identifying determinants of specificity. *Trends Biochem Sci 16*:214–220.

Krause M, Hirsh D. 1987. A *trans*-spliced leader sequence on actin mRNA in *C. elegans*. *Cell 49*:753–761.

Lerner EA, Lerner MR, Janeway CA, Steitz JA. 1981a. Monoclonal antibodies to nucleic acid-containing cellular constituents: Probes for molecular biology and autoimmune disease. *Proc Natl Acad Sci USA 78*:2737–2741.

Lerner MR, Boyle JA, Hardin JA, Steitz JA. 1981b. Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science 211*:400–402.

Maraia RJ, Sasaki-Tozawa N, Driscoll CT, Green ED, Darlington GJ. 1994. The human Y4 small cytoplasmic RNA gene is controlled by upstream elements and resides on chromosome 7 with all other hY scRNA genes. *Nucleic Acids Res 22*:3045–3052.

Matera AG, Frey MF, Margelot K, Wolin SL. 1995. A perinucleolar compartment contains several polymerase III transcripts as well as the polypyrimidine-tract binding protein, hnRNP I. *J Cell Biol*. Forthcoming.

Mattioli M, Reichlin M. 1974. Heterogeneity of RNA protein antigens reactive with sera of patients with systemic lupus erythematosus. *Arthritis Rheumatol 17*:421–429.

Montzka KA, Steitz JA. 1988. Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proc Natl Acad Sci USA 85*:8885–8889.

Nagai K, Oubridge C, Jessen TH, Li J, Evans PR. 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature 348*:515–520.

Nilsen TW. 1993. *Trans*-splicing of nematode premessenger RNA. *Annu Rev Microbiol 47*:413–440.

O'Brien CA, Harley JB. 1990. A subset of hY RNAs is associated with erythrocyte Ro ribonucleoproteins. *EMBO J 9*:3683–3689.

O'Brien CA, Margelot K, Wolin SL. 1993. *Xenopus* Ro ribonucleoproteins: Members of an evolutionarily conserved class of cytoplasmic ribonucleoproteins. *Proc Natl Acad Sci USA 90*:7250–7254.

O'Brien CA, Wolin SL. 1994. A possible role for the 60 kd Ro autoantigen in a discard pathway for defective 5S ribosomal RNA precursors. *Genes & Dev 8*:2891–2903.

Oubridge C, Ito N, Evans PR, Teo CH, Nagai K. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature 372*:432–438.

Palazzolo MJ, Hamilton BA, Ding D, Martin CH, Mead DA, Mierendorf RC, Raghavan KV, Meyerowitz EM, Lipshitz HD. 1990. Phage lambda cDNA cloning vectors for subtractive hybridization, fusion-protein synthesis and Cre-*loxP* automatic plasmid subcloning. *Gene 88*:25–36.

Pruijn GJM, Slobbe RL, van Venrooij WJ. 1991. Analysis of protein–RNA interactions within Ro ribonucleoprotein complexes. *Nucleic Acids Res 19*:5173–5180.

Pruijn GJM, Wingens PAETM, Peters SLM, Thijssen JPH, van Venrooij WJ. 1993. Ro RNP associated Y RNAs are highly conserved among mammals. *Biochim Biophys Acta 1216*:395–401.

Silberklang M, Gillum AM, RajBhandary UL. 1979. Use of in vitro ^{32}P labeling in the sequence analysis of nonradioactive tRNAs. *Methods Enzymol 59*:58–109.

Sulston J, Hodgkin J. 1988. Methods. In: Wood WB, ed. *The nematode Caenorhabditis elegans*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 587–606.

Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol 100*:64–119.

Thomas J, Lea K, Zucker-Aprison E, Blumenthal T. 1990. The splice-

osomal snRNAs of *Caenorhabditis elegans*. *Nucleic Acids Res 18*:2633–2642.

Thomas JD, Conrad RC, Blumenthal T. 1988. The *C. elegans trans*-spliced leader RNA is bound to Sm and has a trimethylguanosine cap. *Cell 54*:533–539.

Van Doren K, Hirsh D. 1988. *Trans*-spliced leader RNA exists as small nuclear ribonucleoprotein particles in *Caenorhabditis elegans*. *Nature 335*:556–559.

van Gelder CWG, Thijssen JPHM, Klaassen ECJ, Sturchler C, Krol A, van Venrooij WJ, Pruijn GJM. 1994. Common structural features of the Ro RNP associated hY1 and hY5 RNAs. *Nucleic Acids Res 22*:2498–2506.

Waterston R, Martin C, Craxton M, Huynh C, Coulson A, Hillier L, Durbin R, Green P, Shownkeen R, Halloran N, Metzstein M, Hawkins T, Wilson R, Berks Du, Z, Thomas K, Thierry-Mieg J, Sulston J. 1992. A survey of expressed genes in *Caenorhabditis elegans*. *Nature Genet 1*:114–123.

Wolin SL, Steitz JA. 1983. Genes for two small cytoplasmic Ro RNAs are adjacent and appear to be single-copy in the human genome. *Cell 32*:735–744.

Wolin SL, Steitz JA. 1984. The Ro small cytoplasmic ribonucleoproteins: Identification of the antigenic protein and its binding site on the Ro RNAs. *Proc Natl Acad Sci USA 81*:1996–2000.

Yoo CJ, Wolin SL. 1994. La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: A yeast homolog of the La autoantigen is dispensable for growth. *Mol Cell Biol 14*:5412–5424.