Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 8640–8644, November 1989
Biochemistry

A second trans-spliced RNA leader sequence in the nematode *Caenorhabditis elegans*

XIN-YUN HUANG AND DAVID HIRSH

Department of Developmental Biology, Synergen, Inc., 1885 33rd Street, Boulder, CO 80301

Communicated by Joan A. Steitz, August 4, 1989

ABSTRACT **In the nematode *Caenorhabditis elegans*, the 22-nucleotide RNA sequence called the spliced leader (SL) is trans-spliced from the 100-nucleotide-long SL RNA to some mRNAs. We have identified a trans-spliced leader (SL2) whose sequence differs from that of the original spliced leader (SL1), although both are 22 nucleotides long. By primer-extension sequencing, SL2 but not SL1 was shown to be present at the 5′ end of the mRNA encoded by one of the four glyceraldehyde-3-phosphate dehydrogenase genes. The other three glyceraldehyde-3-phosphate dehydrogenase genes encode mRNAs that have the SL1 but not the SL2 sequence at their 5′ ends. Therefore, the trans-splicing process can discriminate the transfer of SL1 from that of SL2 in a gene-specific manner.**

The 22 nucleotides (nt) at the 5′ end of three of the four actin mRNAs in *Caenorhabditis elegans* are not encoded contiguously with the protein-encoding portion of the gene (1). This 22-nt untranslated sequence, termed the spliced leader (SL), is identical for all three actin mRNAs. The SL sequence is located in a 1-kilobase (kb) sequence that is tandemly repeated 110 times to form a large array on chromosome V (2, 3). This 1-kb repeat also contains the 5S rRNA gene. This same SL is present on many other mRNAs in addition to actin. It is present in the genomes of all nematodes that have been examined (refs. 4 and 5; S. Bektesh, B. Rosenzweig and D.H., unpublished data).

The joining of the SL to the mRNA occurs through trans-splicing between two independently transcribed precursor RNAs, the SL RNA and the mRNA precursor (6). The SL is derived from a 100-nt SL RNA; the 5′-most 22 nt comprise the SL (1). The SL RNA resembles a typical small nuclear RNA, existing *in vivo* as an anti-Sm antibody-immunoprecipitable SL small nuclear ribonucleoprotein and possessing a trimethylguanosine cap (7–9).

Trans-splicing was first described in trypanosomatid protozoans (10–12). In *Trypanosoma brucei*, a 39-nt SL, derived from a 140-nt SL RNA, is trans-spliced to all mRNAs. The 1.4-kb repeat unit, which contains the 140-nt SL RNA gene, is present in ≈200 tandem copies. The SL RNA in trypanosomes has an unusual modified 5′ terminus with a 7-methylguanosine cap plus four additional modified $O^{2'}$-methyl nucleotides (13, 14). The SL RNA transfers this cap structure to the trans-spliced mRNAs. Unlike trypanosomes, in which the 39-nt spliced leader sequence varies in different species and genera, the 22-nt SL is completely conserved in related species and genera of nematodes (refs. 4 and 15; S. Bektesh, B. Rosenzweig, and D.H., unpublished data). However, within a given species of trypanosome, only one SL is present and it is found on all mRNAs (15). We report here that *C. elegans* contains more than one SL. A second *C. elegans* SL has been found and designated SL2.*

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Primer-Extension Sequencing of RNA. Total RNA was isolated as described (16). Primer-extension sequencing on RNA was done as described by Bektesh *et al.* (4). Oligonucleotide primer (5 ng) specifically labeled with [γ-^{32}P]ATP was annealed to 15 μg of total RNA for 45 min, 2–4°C below the melting temperature of the oligonucleotide. Extension reactions were carried out at 50° for 45 min with avian myeloblastosis virus reverse transcriptase (Life Sciences, Saint Petersburg, FL). Products were separated by electrophoresis in 8% polyacrylamide/7.8 M urea gels and detected by autoradiography.

Isolation of SL2 RNA Genes and DNA Dideoxynucleotide Sequencing. The *C. elegans* EMBL4 genomic library was a gift from Chris Link (University of Colorado, Boulder). The library consists of 15-kb DNA fragments from partial digestion with *Mbo* I, cloned into the λ vector EMBL4. A 20-mer complementary to SL2 was used as the screening probe (see Fig. 3*A*). Dideoxynucleotide sequencing of double-stranded phage λ DNA was done as described by Zaug *et al.* (17).

Northern Blot Analysis. RNAs were either separated in 1.5% agarose/2.2 M formaldehyde gels and transferred to nitrocellulose membrane (Bio-Rad) in 20× SSC (1× SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7) or separated in 5% polyacrylamide/7.8 M urea gels and electroblotted to Hybond-N membrane (Amersham) for 1 hr at 30 V. Oligonucleotide probes were phosphorylated with T4 polynucleotide kinase (Boehringer). Hybridizations were done in 6× SSC/2× Denhardt's solution (1× Denhardt's solution is 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin)/1% SDS with 100 μg of sheared salmon sperm DNA per ml overnight at a temperature 2–4°C below the melting temperature of the oligonucleotide. Washes were done twice at room temperature (15 min each) and at the hybridization temperature for 1 hr in 2× SSC/0.2% SDS.

RESULTS

Identification of a Second trans-Spliced Leader. We found a second trans-spliced leader while determining the 5′ ends of *C. elegans* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNAs by primer-extension sequencing. There are four GAPDH genes in *C. elegans* (18, 19). The sequences of genes *gpd-1* and *gpd-4*, which are on chromosome II, are very similar. They encode the isoenzyme GAPDH-1, which is present in all cells. Genes *gpd-2* and *gpd-3* are tandem direct repeats separated by 244 base pairs (bp) on the X chromosome; they encode the isoenzyme GAPDH-2, which is located in the nematode body wall muscle (refs. 19 and 20; X.-Y.H. and R. Hecht, unpublished data) (Fig. 1*A*). Primer-extension sequencing showed that the previously identified SL is present on mRNAs from three of the four GAPDH

Abbreviations: SL, spliced leader; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; nt, nucleotide(s).

*The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M27263 and M27264).

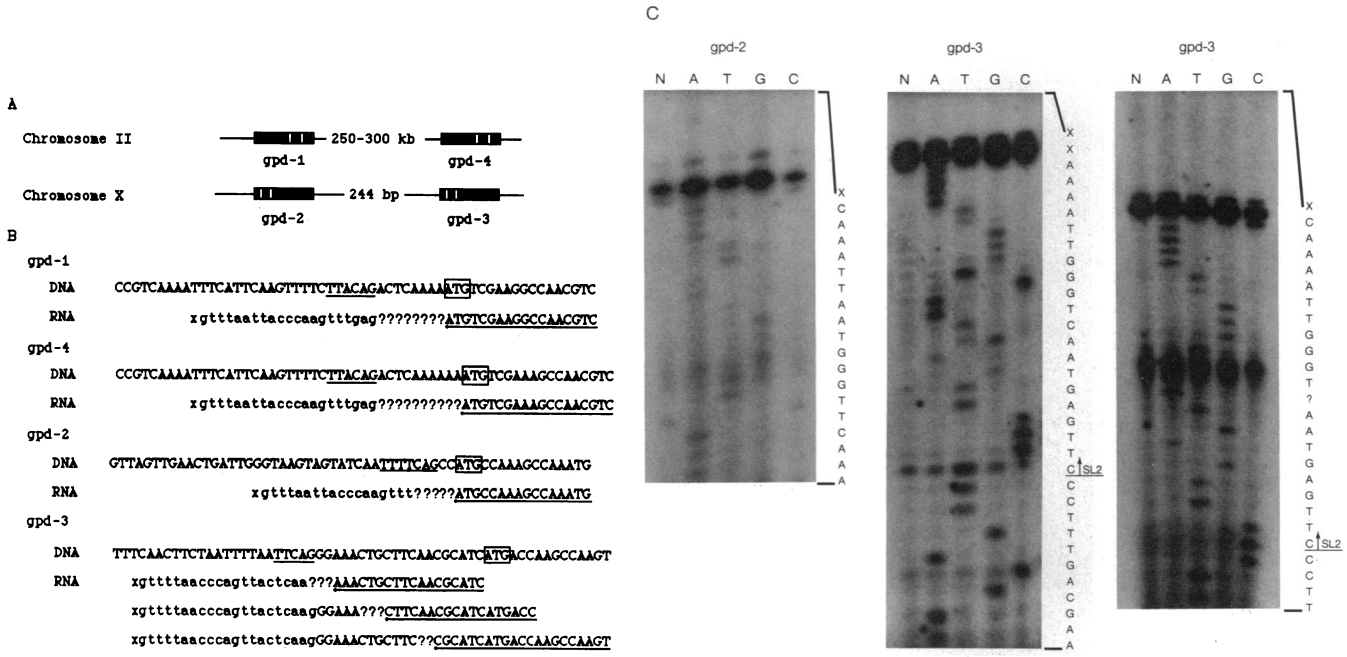

FIG. 1. (*A*) Genomic organization of the GAPDH genes. Black boxes represent exons, and white bars, introns. Transcription is from left to right. (*B*) Primer-extension sequencing of GAPDH mRNAs. Nucleotide sequences of the 5′ ends of GAPDH genes and their mRNAs are compared. The oligonucleotide primers used for extension are underlined with arrows. Genomic 3′ intron splice acceptor sequences are underlined. Start codons are boxed. Lowercase letters indicate the SL sequences. Question mark indicates uncertainty in the sequencing. (*C*) *gpd-2* and *gpd-3* 5′ mRNA sequencing gels. Two *gpd-3* sequencing gels with different primers are shown. Lane N, extension in the absence of dideoxynucleotide. X, unknown first nucleotide.

genes (*gpd-1*, *gpd-2*, and *gpd-4*) (Fig. 1*B*). However, the 22 nt at the 5′ end of *gpd-3* mRNA differ from the corresponding genomic sequence and from the previously described SL (Fig. 1 *B* and *C*). These 22 nt could not be found in the intergenic sequence between *gpd-2* and *gpd-3*, the sequence upstream of *gpd-2*, or the sequence downstream of *gpd-3*.

The point at which the genomic DNA and the mRNA sequences of *gpd-3* diverge corresponds to a consensus 3′ splice acceptor sequence (21) (Fig. 1*B*). This suggests that the new 22 nt is added through RNA splicing. Seventeen of 22 nt are the same in the best alignment of the new 22-nt sequence and the previously described SL sequence (see Fig. 3*A*). This further suggests that the new 22-nt sequence may be a second trans-spliced leader. We designated the previous SL as SL1 and the new one, SL2.

The trans-splicing reaction appears to be gene-specific. The SL1 sequence does not appear in the primer-extension sequencing gel of *gpd-3*, nor does SL2 appear in the sequencing gel of *gpd-2*. [The three cytosines (C9, C10, and C11) near the 5′ end of SL2 serve as a convenient marker for identifying SL2, as do the two uridines (U8 and U9) near the 5′ end for SL1. Neither of these hallmarks can be found in the other's sequencing gel.] These conclusions are limited by the detection of bands in the sequencing gels.

SL2 RNA Gene Organization. SL1 is derived from the 5′ end of the SL1 RNA, which is ≈100 nt long. The SL1 RNA gene and the 5S rRNA gene are present in a 1-kb fragment that is tandemly repeated 110 times in the *C. elegans* genome (1, 2). Association of SL1 RNA genes with the 5S rRNA locus is not required for SL1 RNA gene expression, as the expressed *Haemonchus contortus* SL1 RNA genes are not linked to 5S rRNA genes (S. Bektesh, B. Rosenzweig, and D.H., unpublished data).

Southern blots were used to examine the genomic distribution of SL2 RNA genes in *C. elegans*. An oligonucleotide complementary to 20 of 22 nt within the SL2 was used as the hybridization probe (SL2 probe; see Fig. 3*A*). There appear to be four copies of SL2 RNA genes based on the hybridizing DNA fragments in various restriction enzyme digests (Fig. 2*A*). This analysis indicates that the genomic organization of SL2 RNA genes differs from that of SL1 RNA genes. SL2 RNA genes were isolated and characterized further.

A *C. elegans* EMBL4 genomic library was screened with the SL2 probe in order to isolate the SL2 RNA genes. DNA was prepared from four of the hybridizing phages. These four DNA preparations revealed two different restriction digestion and hybridization patterns. DNA from two isolates overlapped and contained one fragment that hybridized to the SL2 probe. The other two isolates overlapped and contained a single, different, hybridizing fragment. The restriction maps of the two representative phage are shown in Fig. 2*B*. The two classes of genes have been designated SL2 RNA α (represented by phage λSL2/9) and SL2 RNA β (represented by phage λSL2/4). The SL2 RNA β gene has been tentatively mapped by contig mapping (i.e., mapping using a set of overlapping clones) to the region of the major sperm protein gene cluster *msp-142* on chromosome II (A. Coulson and J. Sulston, personal communication). Although the SL2 RNA α gene has not been mapped to a chromosome, it is apparent from the contig mapping that it cannot be within 100–150 kb of *gpd-3* (A. Coulson and J. Sulston, personal communication). Other phage containing potential SL2 RNA genes were not analyzed further.

The nucleotide sequences of SL2 RNA α and β genes and their immediate flanking regions were determined (Fig. 3*A*). The SL2 RNA α and β genes differ by only 7 bases (out of 100) in their transcribed regions. SL2 RNA can be folded into a secondary structure similar to that proposed for the original SL1 RNA (8) (Fig. 3*B*). The genes share a high level of sequence homology near the 5′ region and less homology in the 3′ region. SL2 RNA genes show similarity to the SL1 RNA gene sequence at several critical positions, including the splice donor site and consensus Sm-binding sequences (Fig. 3*A*). The conserved sequences at the 5′ flanking region, especially the long match from −20 to −70 in Fig. 3*A*, may be important for the regulation of SL RNA gene expression.

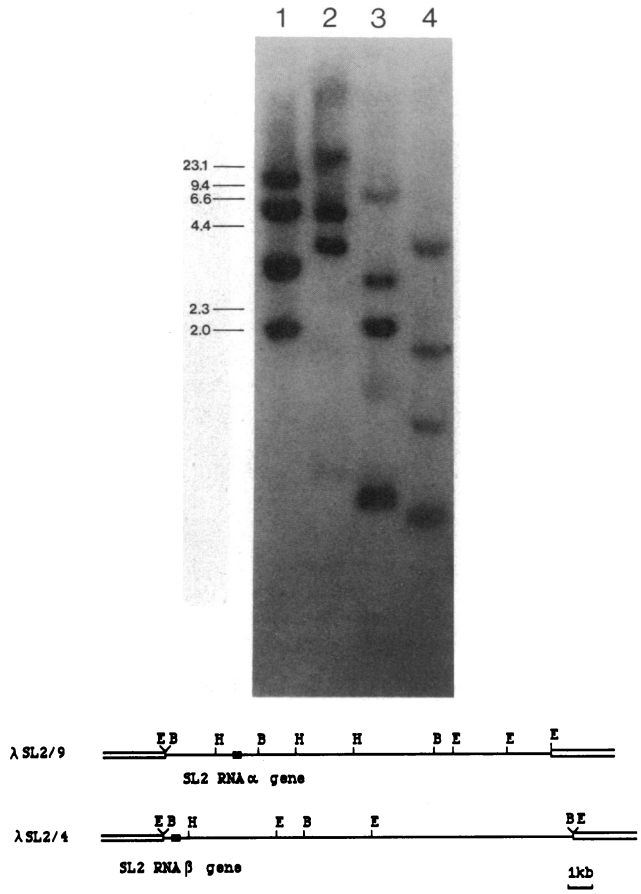

FIG. 2. (*A*) Genomic Southern blot. *C. elegans* genomic DNA was cleaved with various restriction enzymes, electrophoresed (10 μg per lane) in an agarose gel, transferred to nitrocellulose, and hybridized with the SL2 probe. Indicated molecular sizes (kb) were determined with λ *Hin*dIII markers. Lane 1, *Eco*RI; lane 2, *Bam*HI; lane 3, *Hin*dIII; lane 4, *Hpa* II. (*B*) Restriction maps of two phages containing SL2 RNA genes. Phage λSL2/9 contains the SL2 RNA α gene, and λSL2/4 contains the SL2 RNA β gene. Solid boxes mark the fragments encoding SL2 RNA sequences. E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII.

Presence of SL2 on Other *C. elegans* mRNAs and in Other Nematodes. SL1 is present in all nematodes that have been analyzed but it has not been found in other eukaryotes (4, 5). An oligonucleotide complementary to SL1 arrested the translation of ≈10% of the proteins visible on a two-dimensional gel after *in vitro* translation of *C. elegans* mRNAs in a rabbit reticulocyte system (4). Therefore, ≈10% of the *C. elegans* mRNAs appear to acquire SL1. Northern analyses were used to examine whether other RNAs in addition to *gpd-3* mRNA contain SL2 and to determine whether SL2 exists in other nematodes (Fig. 4). Total RNA was hybridized to the ^{32}P-labeled SL2 probe. In *C. elegans* (lane 1), a smear of RNAs representing a wide range of sizes was detected. These RNAs probably represent transcripts derived from genes other than *gpd-3* that also contain the SL2 sequence. SL2 is also present in RNAs isolated from *C. elegans* var. Bergerac (lane 2) and *C. briggsae* (lane 3), but not in RNAs from the nematodes *P. redivivus* (lane 4) and *H. contortus* (lane 5). In this respect, SL2 differs from SL1, which is found in mRNAs in all nematodes (4, 5). SL2 is not found in *Dictyostelium* or human RNAs (data not shown). Although *C. briggsae* has the homologous *gpd-3* gene, *C. elegans* var. Bergerac does not; therefore, other RNAs in *C. elegans* var. Bergerac must have SL2, which corroborates the observation in Fig. 4 (ref. 19; Y. L. Lee, X.-Y.H., and R. Hecht, unpublished data).

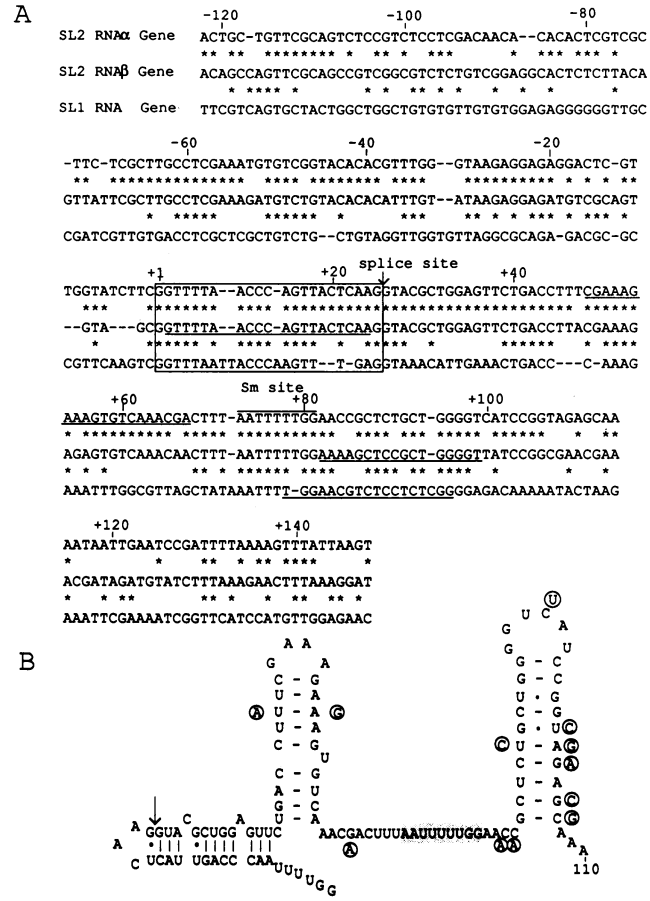

FIG. 3. (*A*) Nucleotide sequences of SL2 RNA α and β genes compared with that of the SL1 RNA gene. The nucleotide sequence of the sense strand is presented with numbering from the 5′ end of the 22-nt SL, which is boxed. Underlined segments represent sequences for which complementary oligonucleotides were synthesized and used in subsequent experiments. The splice donor site and Sm binding site are marked. For sequence comparisons, gaps (dashes) were introduced in the sequences to maximize similarity. Numbering of the nucleotides includes the gaps. (*B*) Proposed secondary structure of SL2 RNAs. Arrow indicates the 5′ splice site. Circled nucleotides are substitutes in SL2 RNA β. Stippled area is the Sm site. The exact 3′ ends of both SL2 RNA transcripts are not known. The Δ*G* value of this structure is −28.4 kcal/mol (1 kcal = 4184 J). The program used was from the University of Wisconsin Genetics Computer Group.

SL2 Is Derived from Either a 110- or a 100-nt Precursor RNA. Northern blots were used to detect the RNA transcripts from which the SL2 is derived. Total RNAs from *C. elegans* var. Bristol, *C. elegans* var. Bergerac, and *C. briggsae* were resolved by polyacrylamide gel electrophoresis in denaturing conditions and hybridized to a labeled synthetic oligonucleotide complementary to positions +48 to +67 in the SL2 RNA α gene (Fig. 3*A*). With the α probe, a single band was detected in the lanes containing *C. elegans* RNA and no signal was observed in the lane containing RNA from *C. briggsae*. The band is at 110 nt based on the markers used (Fig. 5*A*). Stringent hybridization conditions (2°C below the melting temperature of the oligonucleotide) were used because this 20-mer (α probe) has only two bases different from the same region in SL2 RNA β. At lower stringency, the α probe cross-hybridized with SL2 RNA β (data not shown). Hybridizing the filter with a 17-mer complementary to positions +82 to +99 in the SL2 RNA β gene (β probe) revealed a 100-nt band in *C. elegans* and *C. briggsae* RNA (Fig. 5*B*). These results show that both SL2 genes are transcribed. SL2

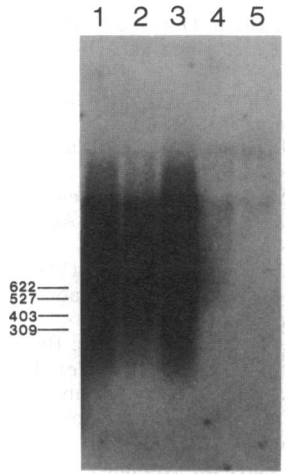

FIG. 4. Presence of SL2 in multiple RNAs in *C. elegans* and *C. briggsae* shown by Northern blot analyses. Ten micrograms of total nematode RNA was fractionated and hybridized to the oligonucleotide complementary to SL2. Indicated sizes (nt) were determined by using DNA size standards. Lane 1, *C. elegans* var. Bristol; lane 2, *C. elegans* var. Bergerac; lane 3, *Caenorhabditis briggsae*; lane 4, *Panagrellus redivivus*; lane 5, *Haemonchus contortus*.

RNA α exists only in *C. elegans* and SL2 RNA β is present in *C. elegans* and *C. briggsae*. The RNA blot was also probed with an 18-mer (+78 to +96) specific for SL1 RNA (Fig. 5*C*). The result is consistent with SL1 RNA being ≈100 nt long and identical in *C. elegans* and *C. briggsae* (ref. 1; D. W. Nelson and B. M. Honda, personal communication).

DISCUSSION

The results demonstrate that the 22-nt sequence at the 5′ end of *gpd-3* mRNA is a trans-spliced leader other than SL1. Two SL2 RNA genes (α and β) of *C. elegans* have been isolated and sequenced. We have identified RNAs of 110 or 100 nt as their transcription products.

The organization of the SL1 and SL2 RNA genes in the *C. elegans* genome are different from each other. The SL1 RNA genes are present in a 1-kb tandem repeat unit of 110 copies per haploid genome. In contrast, each of the four SL2 RNA genes is apparently present at a different locus in the genome; the two SL2 RNA genes differ in their coding regions by 7 base substitutions out of 100 bases. Both kinds of gene organization and copy numbers have been seen in genes encoding U-type small nuclear RNAs (see ref. 22 for review).

Finding a second spliced leader in *C. elegans* raises several questions. It has been shown that *gpd-2* and *gpd-3* mRNAs are body wall muscle-specific (ref. 19; X.-Y.H. and R. Hecht, unpublished data). Genes *gpd-2* and *gpd-3* are direct tandem repeats and encode the same isoenzyme. However, their mRNAs acquire different SLs at their 5′ ends. Since more than one SL exists in the same cells (i.e., the body wall muscle cells), the question arises as to how a particular SL is trans-spliced only to a specific pre-mRNA. It is reasonable to suppose that the 5′ untranslated region is sufficient for recognition by the trans-splicing machinery, because *gpd-2* and *gpd-3* coding regions are so similar (19). Since the 3′ splice acceptor site can be adjacent to the initiation codon, the untranslated sequence 5′ to the 3′ splice site and the 3′ splice site itself might be enough for the specific selection of the SL if the coding sequence is not utilized (4). In cis-splicing the branch point and the 3′ splice acceptor site are important for 3′ splice-site selection (23, 24). From the comparison of SL1 and SL2 RNAs, it is also possible that the conserved regions and/or the surrounding nonconserved regions in SL1 and SL2 RNAs may be involved in specifying which SL is trans-spliced. It is possible that specific proteins associated with the individual SL RNAs play an important role in this specification process.

The functional significance of the trans-splicing process and the SL itself remains unknown. In trypanosomes, it has been suggested that trans-splicing is involved in the 5′ end processing of polycistronic precursors into monocistronic mature mRNAs (25). Perhaps there is a preference for only a short sequence of nucleotides immediately before the initiation codon to accelerate ribosomal scanning. Therefore, one of the possible functions of trans-splicing might be to remove the long 5′ untranslated region and other upstream potential start codons. Additional functions or complex interactions for SL1 are likely to exist, since the 22-nt sequence of SL1 is conserved throughout widely divergent nematodes (4, 5). This conservation could be important for protein-binding in SL ribonucleoprotein formation or for translational control.

The sequence homology and similarity in secondary structure of SL1 and SL2 RNAs suggest that these RNAs may

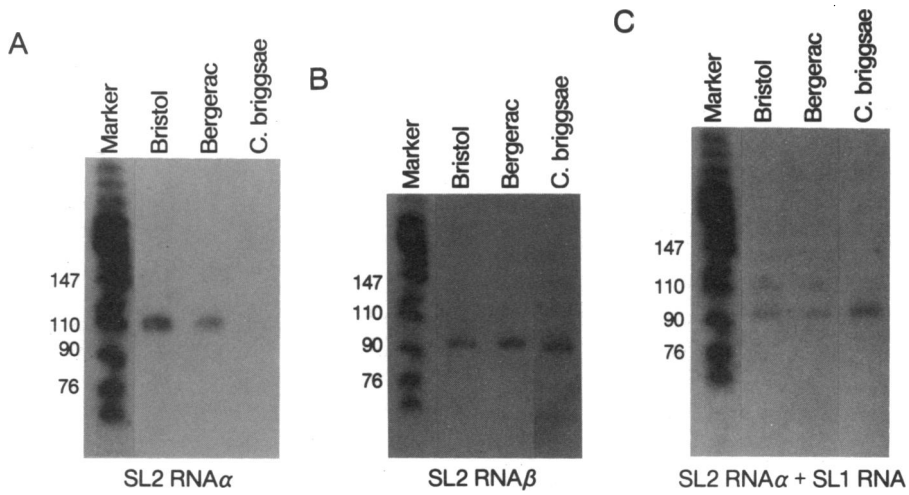

FIG. 5. Detection of the SL2 RNA. (*A*) Autoradiograph of Northern blot of total nematode RNA (10 μg per lane) probed with the labeled 20-mer complementary to positions +48 to +67 in SL2 RNA α gene. Markers were labeled products of an *Msp* I digest of pBR322. RNA samples were from *C. elegans* var. Bristol, *C. elegans* var. Bergerac, and *C. briggsae*. (*B*) Same blot, probed with a 17-mer complementary to positions +82 to +99 in SL2 RNA β gene. (*C*) Same blot, probed with an oligonucleotide (+78 to +96) specific for the *C. elegans* SL1 RNA. The probe for SL2 RNA α was not stripped off before probing for SL1.

serve similar functions and that they may be derived from a common evolutionary precursor. Both SL1 and SL2 RNAs contain trimethylguanosine caps that are transferred to and maintained on the trans-spliced mRNAs (K. Van Doren and D.H., unpublished results; R.-F. Liou, J. D. Thomas, and T. Blumenthal, personal communication; X.-Y.H. and D.H., unpublished results). The reason why *gpd-2* acquires SL1 and *gpd-3* has SL2 is not clear. The data described here show that both SL2 RNA α and β genes are transcribed and that both are capable of donating the same 22-nt SL2. The levels of SL2 RNA α and β are lower than those of SL1 RNA (unpublished results). This could be due to a higher number of copies of the SL1 RNA genes. The results presented in Fig. 3 show that short sequences are conserved between the 5′ flanking regions of SL1 and SL2 RNA genes. These sequences might modulate expression of the genes.

A detailed analysis of trans-splicing in *C. elegans* would be aided by information on which individual nucleotides within the SL and the pre-mRNA are essential. Mutagenesis and transformation studies should identify which sequences in the molecules are important in trans-splicing.

We are grateful to Ralph Hecht and Susan Bektesh for their involvement in the initiation of this work. We thank Jim Bruzik for help with the secondary structure determination of SL2 RNA and Alan Coulson and John Sulston for the contig mapping of SL2 RNA genes. We thank our colleagues for helpful discussions and Kevin Van Doren, Joe Cox, and Mike Milhausen for critical reading of the manuscript. This investigation was supported by Public Health Service Grant GM37823.

1. Krause, M. & Hirsh, D. (1987) *Cell* **49**, 753–761.
2. Nelson, D. W. & Honda, B. M. (1985) *Gene* **38**, 245–251.
3. Albertson, D. G. (1985) *EMBO J.* **4**, 2493–2498.
4. Bektesh, S., Van Doren, K. & Hirsh, D. (1988) *Genes Dev.* **2**, 1277–1283.
5. Takacs, A. M., Denker, J. A., Perrine, K. G., Moroney, P. A. & Nilsen, T. W. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7932–7936.
6. Bektesh, S. & Hirsh, D. (1988) *Nucleic Acids Res.* **16**, 5692.
7. Van Doren, K. & Hirsh, D. (1988) *Nature (London)* **335**, 556–559.
8. Bruzik, J. P., Van Doren, K., Hirsh, D. & Steitz, J. A. (1988) *Nature (London)* **335**, 559–562.
9. Thomas, J. D., Conrad, R. C. & Blumenthal, T. (1988) *Cell* **54**, 533–539.
10. Borst, P. (1986) *Annu. Rev. Biochem.* **55**, 701–732.
11. Murphy, W. J., Watkins, K. P. & Agabian, N. (1986) *Cell* **47**, 517–525.
12. Sutton, R. E. & Boothroyd, J. C. (1986) *Cell* **47**, 527–535.
13. Perry, K. L., Watkins, K. P. & Agabian, N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8190–8194.
14. Freistadt, M. S., Cross, G. A. M., Branch, A. D. & Robertson, H. D. (1987) *Nucleic Acids Res.* **15**, 9861–9879.
15. Walder, J. A., Eder, P. S., Engman, D. M., Brentano, S. T., Walder, R. Y., Knutzon, D. S., Dorfman, D. M. & Donelson, J. E. (1986) *Science* **233**, 569–571.
16. Cox, G. N., Carr, S., Krammer, J. M. & Hirsh, D. (1985) *Genetics* **109**, 513–528.
17. Zaug, A. J., Kent, J. R. & Cech, T. R. (1984) *Science* **224**, 574–578.
18. Yarbrough, P. O., Hayden, M. A., Dunn, L. A., Vermersch, P. S., Klass, M. R. & Hecht, R. M. (1987) *Biochim. Biophys. Acta* **908**, 21–33.
19. Huang, X.-Y., Barrios, L. A. M., Vonkhorporn, P., Honda, S., Albertson, D. G. & Hecht, R. M. (1989) *J. Mol. Biol.* **206**, 411–424.
20. Yarbrough, P. O. & Hecht, R. M. (1984) *J. Biol. Chem.* **259**, 14711–14720.
21. Emmons, S. W. (1987) in *The Nematode Caenorhabditis elegans*, ed. Wood, W. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 47–79.
22. Dahlberg, J. E. & Lund, E. (1988) in *Small Nuclear Ribonucleo-protein Particles*, ed. Birnstiel, M. L. (Springer, New York), pp. 38–70.
23. Reed, R. & Maniatis, T. (1986) *Cell* **46**, 681–690.
24. Reed, R. & Maniatis, T. (1988) *Genes Dev.* **2**, 1268–1276.
25. Boothroyd, J. C. (1989) in *Nucleic Acids and Molecular Biology*, eds. Eckstein, F. & Lilley, D. M. J. (Springer, Berlin), Vol. 3, in press.