MOLECULAR AND CELLULAR BIOLOGY, Aug. 1989, p. 3543–3547
0270-7306/89/083543-05$02.00/0

Vol. 9, No. 8

Characterization and Expression of a Spliced Leader RNA in the Parasitic Nematode *Ascaris lumbricoides* var. *suum*

TIMOTHY W. NILSEN,[1]* JOSEPH SHAMBAUGH,[1] JOHN DENKER,[1] GRETCHEN CHUBB,[1] CHRISTINE FASER,[2] LAURA PUTNAM,[2] AND KAREN BENNETT[2]

Department of Molecular Biology and Microbiology, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106,[1] and Department of Microbiology, School of Medicine, University of Missouri, Columbia, Missouri 65212[2]

Received 12 January 1989/Accepted 3 May 1989

The parasitic nematode *Ascaris* spp. contains a 22-nucleotide spliced-leader (SL) sequence identical to the *trans*-SL previously described in *Caenorhabditis elegans* and other nematodes. The SL comprises the first 22 nucleotides of a ~110-base RNA and is transcribed by RNA polymerase II. The SL RNA contains a trimethylguanosine cap and a consensus Sm binding site. Furthermore, the *Ascaris* SL RNA has the potential to adopt a secondary structure which is nearly identical to potential secondary structures of similar SL RNAs in *C. elegans* and *Brugia malayi*.

A subset of mRNAs in the nematode *Caenorhabditis elegans* contains 22-nucleotide spliced-leader (SL) sequences at the 5′ ends (2, 9). Several lines of evidence, including the identification of appropriate branched intermediates, indicate that the SL is acquired through an intermolecular (*trans*) splicing reaction (1, 9). In this reaction, the SL is donated from the 5′ end of an ~100-base RNA (SL RNA) transcribed from within the 5S rRNA gene cluster (9). The SL RNA of *C. elegans* has many features in common with well-characterized small nuclear RNAs (snRNAs) known to be essential for normal *cis* splicing. The SL RNA contains a trimethylguanosine ($m_3^{2,2,7}G$) cap and has a consensus binding site (RAU_nGR) for Sm, an antigen associated with small nuclear ribonucleoproteins (4, 5, 14, 15). This Sm binding site is functional, since SL RNA is immunoprecipitated by anti-Sm antisera from extracts of *C. elegans* or after incubation in HeLa cell nuclear extracts (5, 14, 15). Furthermore, the sequence elements presumed to mediate transcriptional initiation and termination of SL RNA resemble those of vertebrate snRNAs (5). These properties suggest that the SL RNA might have a dual function in the *trans*-splicing process in which the 5′ donated exon is covalently linked to an snRNA-like sequence (5, 15).

trans splicing in nematodes does not appear to be restricted to *C. elegans*. Some mRNAs in the filarial parasite *Brugia malayi* contain a 22-base SL identical in sequence to the *C. elegans* SL (13). In *B. malayi*, as in *C. elegans*, the SL is acquired from an SL RNA transcribed from within the 5S rRNA gene cluster (13). Nucleotide sequence analysis of genomic clones has established the presence of the same 22-nucleotide SL sequence in *Panagrellis redivivus* and *Haemonchus contortus*, and blot hybridization analysis has suggested that the same SL sequence may be present on mRNAs of *Ascaris* and *Anisakis* spp. (2). On the basis of these observations, it seems likely that *trans* splicing involving acquisition of the 22-nucleotide SL is widespread, if not universal, in nematodes.

Because of its numerous advantages for biochemical and developmental studies, we have characterized the potential SL RNA of *Ascaris* spp. We show that the 22-nucleotide SL in this organism is identical in sequence to those in other nematodes. A ~110-base SL RNA is transcribed from within the *Ascaris* 5S rRNA gene cluster, possesses a trimethylguanosine cap, and contains a consensus Sm binding site. The SL RNA is transcribed by RNA polymerase II and is expressed during early embryonic development. Other than the 22-nucleotide SL sequence itself, there is little primary sequence conservation between the SL RNA of *Ascaris* spp. and SL RNAs of *C. elegans* and *B. malayi*. However, all three SL RNAs can potentially adopt similar secondary structures.

To examine the organization of SL-related sequences in *Ascaris* spp., an oligonucleotide complementary to the 22-nucleotide SL was used to probe Southern blots of genomic DNA digested with a variety of restriction enzymes. Parallel blots were also probed with an oligonucleotide complementary to *B. malayi* 5S rRNA, since in six of the seven nematodes so far examined, the SL sequence is encoded within the 5S rRNA locus (2, 13). In DNA digested with *Sca*I, *Cla*I, or *Hae*III, both probes identified a reiterated 1-kilobase fragment (data not shown). This result suggested that the SL-related sequence was linked to the 5S rRNA sequence in *Ascaris* spp. To characterize the SL sequence in detail, an *Ascaris* genomic library (in λ EMBL4) (3) was screened by hybridization with both the SL and 5S oligonucleotides. Plaques which hybridized to both probes were selected and purified by three successive rounds of replating and hybridization. Restriction enzyme mapping of isolated bacteriophage DNA indicated that the genomic inserts contained a tandemly reiterated 1-kilobase fragment. The nucleotide sequence of this fragment is shown in Fig. 1. A 22-nucleotide SL sequence identical to that described in other nematodes was present 320 nucleotides downstream of the 3′ end of a gene encoding 5S rRNA and was in the same transcriptional orientation as the 5S RNA. Additionally, the 22-nucleotide SL was immediately flanked on the 3′ side by the dinucleotide GT characteristic of a potential splice donor site, and a consensus Sm binding site, AAUUUUGG, was found 53 nucleotides downstream of the SL sequence. The organization of the SL sequence and 5S rRNA sequence in *Ascaris* spp. is the same as that in *B. malayi* and contrasts with the organization of these sequences in *C. elegans*, in which the SL is transcribed in the opposite orientation to 5S rRNA (9).

Preliminary Northern (RNA) blot analysis of *Ascaris* nonpolyadenylated RNA derived from various adult tissues

* Corresponding author.

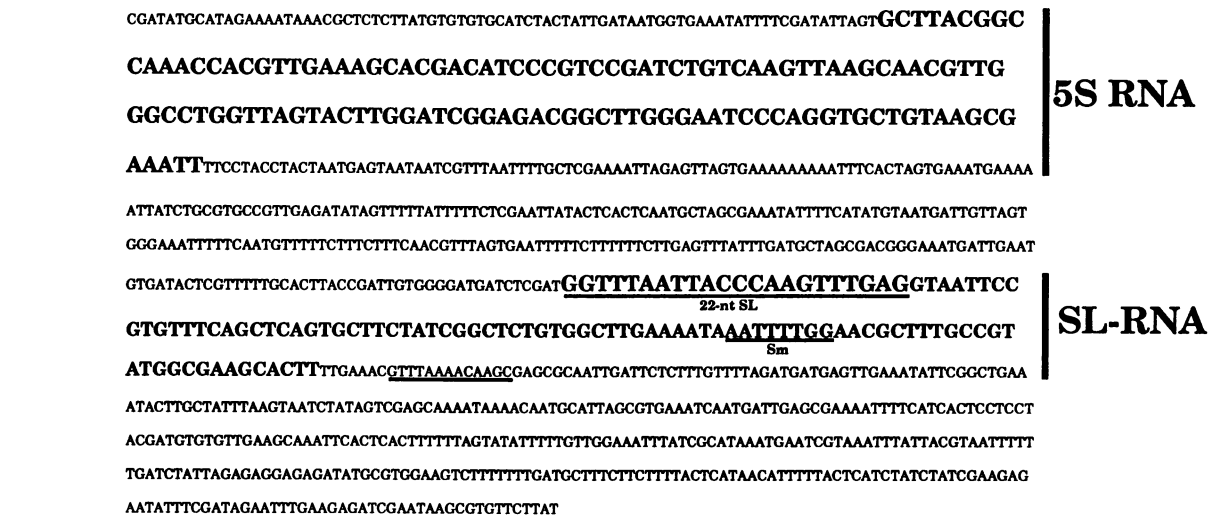

FIG. 1. Nucleotide sequence of the SL and 5S rRNA gene repeats in *Ascaris* spp. A 1-kilobase repeated *Cla*I fragment from a genomic clone containing *Ascaris* SL RNA and 5S rRNA coding sequences (see text) was subcloned into pBSm13+ (Stratagene) and sequenced by the dideoxynucleotide chain terminator method (12). 5S RNA and SL RNA coding regions are indicated. The 22-nucleotide (nt) SL, a potential Sm binding site, and a potential termination site for SL RNA transcription, are underlined.

or stages of embryonic development by using an oligonucleotide probe complementary to the SL sequence indicated the presence of a discrete RNA of approximately 110 nucleotides (data not shown). To determine if this RNA was transcribed from within the 5S locus, an oligonucleotide complementary to bases 33 through 54 downstream of the SL sequence was constructed and used for both Northern blot and primer extension analyses. This oligonucleotide identified a ~110-nucleotide RNA on Northern blots. Primer extension sequencing using this oligodeoxynucleotide indicated that the ~110-nucleotide RNA contained the 22-nucleotide SL sequence at its 5′ end (data not shown). We concluded from these experiments that in *Ascaris* spp., a ~110-base SL RNA is transcribed from within the 5S rRNA gene repeat. When sequences encoding the SL RNAs of *Ascaris* spp., *C. elegans*, and *B. malayi* were compared, identity was confined to the 22-nucleotide SL sequence, the potential splice donor sites, and the potential Sm binding sequences. However, computer-generated secondary structures of these three RNAs were remarkably similar (Fig. 2).

To assess the cap structure present on the SL RNA, total nonpolyadenylated *Ascaris* RNA was 3′ end labeled by using [^{32}P]pCp and RNA ligase (6) and immunoprecipitated with a monoclonal antibody specific for trimethylguanosine caps (generously provided by Adrian Krainer, Cold Spring Harbor Laboratory; 8). The same analysis was performed on *B. malayi* RNA. A similar but nonidentical spectrum of RNAs was precipitated in the two organisms (Fig. 3). To determine which, if any, of the bands corresponded to SL RNAs, the immunoprecipitated RNAs were annealed with the oligonucleotide complementary to the SL sequence and digested with RNase H. In both cases, this treatment resulted in the shift in size of a specific precipitated RNA (Fig. 3). This result suggested that the SL RNAs of *Ascaris* spp. and *B. malayi* contained a trimethylguanosine cap. To confirm this interpretation, the end-labeled RNAs identified by RNase H treatment were excised from preparative gels and subjected to direct sequence analysis by partial chemical degradation (11). The obtained sequence confirmed the identity of the immunoprecipitated RNAs with the SL RNAs of these two organisms (data not shown).

To determine the nature of the RNA polymerase activity responsible for SL RNA synthesis, nuclear run-on experiments were performed with or without the addition of the RNA polymerase inhibitor α-amanitin (1 μg/ml). For this experiment, nuclei were prepared from synchronous 60-cell (gastrula) embryos (P. J. Cleavinger, J. W. McDowell, and K. L. Bennett, Dev. Biol., in press). In these nuclei, the SL RNA sequence was actively transcribed, and this transcription was effectively inhibited by addition of 1 μg of α-amanitin per ml (Fig. 4). As expected, transcription of actin sequences was also inhibited by this concentration of α-amanitin, while 26S rRNA transcription was resistant to the inhibitor (Fig. 4). We infer from these results that the SL RNA of *Ascaris* spp. is transcribed by RNA polymerase II. Similar experiments (M. Golomb and S. Becktesh, personal communication) suggest that the SL RNA in *C. elegans* is also transcribed by RNA polymerase II.

We have shown here that *Ascaris* spp. contain a SL sequence identical to that previously described in *C. elegans* and other nematodes. As in *C. elegans* and *B. malayi*, the SL is found at the 5′ end of a small nonpolyadenylated RNA (SL RNA) transcribed from within the 5S rRNA gene repeat.

The results of Northern blot hybridizations suggest that the SL sequence is present on many *Ascaris* mRNAs (2; our unpublished observations). While it seems likely that this SL is derived from the *Ascaris* SL RNA by *trans* splicing, it should be pointed out that direct evidence for *trans* splicing (in the form of predicted intermediates) has been obtained only in *C. elegans* (1).

As noted above, several investigators (4, 5, 14, 15) have documented similarities between the *C. elegans* SL RNA and snRNAs. The *Ascaris* SL RNA shares these similarities, including a trimethylguanosine cap, a potential Sm binding site, and transcriptional signals typical of vertebrate snRNAs. In this regard a consensus termination signal for vertebrate snRNAs ($GTTTN_{0-3}AAA^{G}_{A}NNAGA$) is located 9 to 28 nucleotides downstream of snRNA 3′ ends (7, 10, 17). An almost identical sequence, $GTTTAAAA^{CA}AGC$, appears eight nucleotides downstream of the *Ascaris* SL RNA coding region. While it is tempting to speculate on the potential functions of this and other putative transcriptional control elements, definitive proof of their roles in SL RNA

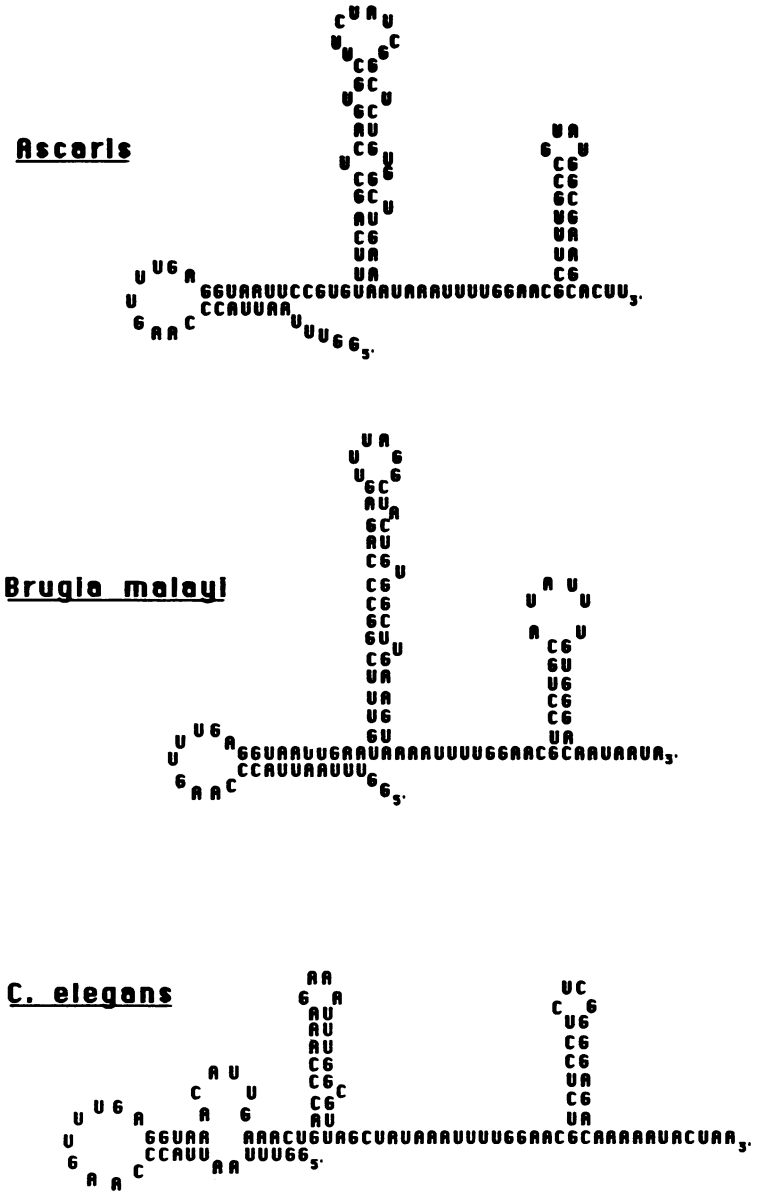

FIG. 2. Potential secondary structures of SL RNAs from *Ascaris* spp., *B. malayi*, and *C. elegans*. The structures were generated by using the computer program of Zuker and Stiegler (18) with the constraint that Sm binding sites were kept unpaired. The *C. elegans* structure is drawn according to Bruzik et al. (5).

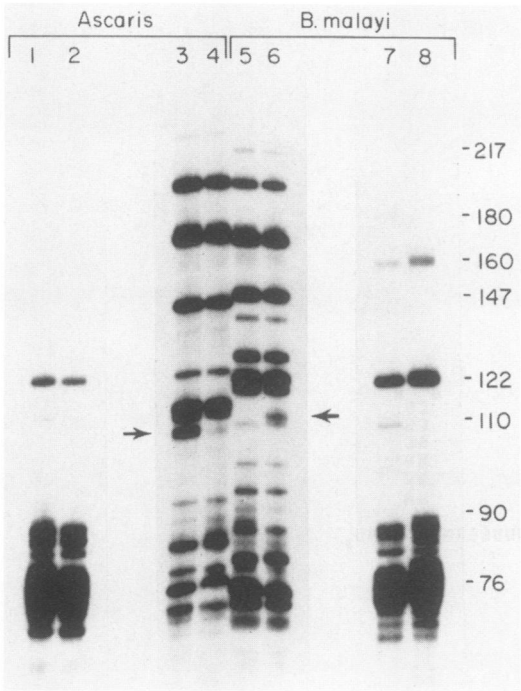

FIG. 3. SL RNAs from *B. malayi* and *Ascaris* spp. contain a trimethylguanosine cap structure. Nonpolyadenylated RNAs from *Ascaris* spp. (lanes 1 through 4) or *B. malayi* (lanes 5 through 8) were 3′ end-labeled with [^{32}P]pCp as described in the text, immunoprecipitated with antitrimethylguanosine antisera as described elsewhere (4) (lanes 3 through 6), and fractionated on a 6% denaturing polyacrylamide gel before autoradiography. Lanes 4 and 5, Precipitated RNAs annealed with 100 ng of an oligonucleotide complementary to the SL sequence and digested with RNase H (16) prior to electrophoresis. The arrows indicate the position of RNAs whose electrophoretic mobilities were altered by this treatment. Lanes 1 and 8, Labeled RNAs prior to immunoprecipitation; lanes 2 and 7, RNAs which were not bound by the antitrimethylguanosine antibody. Lanes 1, 2, 7, and 8 were exposed for 3 h, and lanes 3 through 6 were exposed for 20 h. The indicated sizes correspond to the positions of labeled marker restriction fragments electrophoresed in parallel lanes. Numbers on the right indicate molecular sizes in nucleotides.

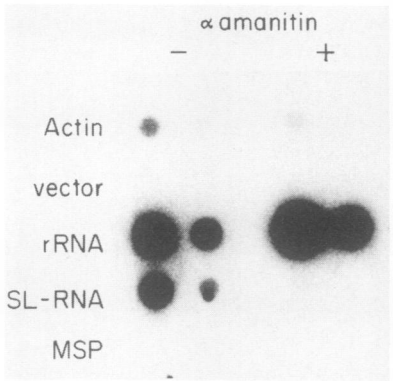

FIG. 4. SL RNA of *Ascaris* spp., as transcribed by RNA polymerase II. Nuclei were prepared from 60-cell *Ascaris* embryos, and nuclear run-on transcriptions in the presence of [^{32}P]UTP were performed as described elsewhere (Cleavinger et al., in press). Following phenol extraction and ethanol precipitation, equal portions of labeled run-on transcription reactions were hybridized to filters with either 1 or 0.1 μg of a 334-base *Ssp*I subclone containing the *Ascaris* SL RNA sequence (derived from the 1-kilobase [Fig. 1] *Ascaris* 5S rRNA repeat) (SL-RNA), an *Ascaris* major sperm protein (MSP) cDNA clone (3), a subclone of *Ascaris* 26S rRNA (rRNA) (Cleavinger et al., in press), an *Ascaris* actin genomic clone (Actin) (Cleavinger et al., in press), and plasmid sequence (vector), as described elsewhere (Cleavinger et al., in press). Major sperm protein, which is transcribed only in the testis (3), and vector sequences served as negative hybridization controls. Reactions were performed in the absence or presence of 1 μg of α-amanitin per ml as indicated. Following hybridization, filters were washed and autoradiographed as described elsewhere (Cleavinger et al., in press).

transcription awaits the development of appropriate in vitro systems.

Bruzik et al. have generated a computer-assisted potential secondary structure of the *C. elegans* SL RNA and have noted a striking similarity between this structure and similarly generated structures of the SL transcripts of various trypanosomatids (5). Each of these structures contains three stem loops with the 5′ splice site adjacent to the turn of the 5′-most loop (5). Bruzik et al. have suggested the possibility that SL small nuclear ribonucleoproteins, like U_1 in conventional *cis* splicing, might autonomously activate their own 5′ splice sites and, in doing so, obviate the necessity for U_1 in *trans*-splicing reactions (5). Despite considerable divergence in the primary sequences, we have shown that the potential secondary structures of two additional nematode SL RNAs are highly conserved and equivalent to the *C. elegans* structure. While this conservation is likely to be important, it should also be noted that the 22-nucleotide sequence itself is completely conserved among these widely divergent nematodes. This conservation, especially in the 5′ loop, is unlikely to result from structural constraints and may reflect a requirement for protein recognition during either processing or some subsequent event in nematode mRNA metabolism. In this regard, the role, if any, of the SL on a specific subset of nematode mRNAs remains to be established.

We are grateful to Adrian Krainer, Cold Spring Harbor Laboratory, for providing antitrimethylguanosine antibody and to Patricia Maroney and Greg Hannon for helpful discussions.

This work was supported by a grant from the John D. and Catherine T. MacArthur Foundation, a P.Y.I. award of the National Science Foundation to T.W.N., and grants from the Rockefeller Foundation and the March of Dimes to K.B.

LITERATURE CITED

1. **Bektesh, S. L., and D. I. Hirsh.** 1988. *C. elegans* mRNAs acquire a spliced leader through a *trans*-splicing mechanism. Nucleic Acids. Res. **16**:5692.
2. **Bektesh, S. L., K. V. van 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.
3. **Bennett, K. L., and S. Ward.** 1986. Several somatically expressed genes are lost or rearranged during embryonic chromatin dimunition in the nematode *Ascaris lumbricoides* var *suum*. Dev. Biol. **118**:141–147.
4. **Blumenthal, T., and J. Thomas.** 1988. *cis* and *trans* mRNA splicing in *C. elegans*. Trends Genet. **4**:305–308.
5. **Bruzik, J. P., K. van Doren, D. Hirsh, and J. A. Steitz.** 1988. *trans*-splicing involves a novel form of small ribonucleoprotein particles. Nature (London) **335**:559–562.
6. **England, T. E., and O. C. Uhlenbeck.** 1980. Specific labeling of 3′-termini of RNA with T4 RNA ligase. Methods Enzymol. **65**:65–74.
7. **Hernandez, N.** 1985. Formation of the 3′ end of U1 snRNA is directed by a conserved sequence located downstream of the coding region. EMBO J. **4**:1827–1837.

8. **Krainer, A. R.** 1988. Pre-mRNA splicing by complementation with purified human U_1, U_2, U_4U_6 and U_5 snRNPs. Nucleic Acids. Res. **16**:9415–9429.
9. **Krause, M., and D. Hirsh.** 1987. A *trans*-spliced leader sequence on actin mRNA in *C. elegans*. Cell **49**:753–761.
10. **Neuman de Vegvar, H. E., E. Lund, and J. E. Dahlberg.** 1986. 3′ end formation of U_1 snRNA precursors is coupled to transcription from snRNA promoters. Cell **47**:259–266.
11. **Peattie, D. A.** 1979. Direct chemical method for sequencing RNA. Proc. Natl. Acad. Sci. USA **76**:1760–1764.
12. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
13. **Takacs, A. M., J. A. Denker, K. G. Perrine, P. A. Maroney, and T. W. Nilsen.** 1988. A 22-nucleotide spliced leader sequence in the human parasitic nematode *Brugia malayi* is identical to the *trans*-spliced leader exon in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **85**:7932–7936.
14. **Thomas, J. D., R. C. Conrad, and T. Blumenthal.** 1988. The *C. elegans trans*-spliced leader RNA is bound to Sm and has a trimethylguanosine cap. Cell **54**:533–539.
15. **Van Doren, K. V., and D. Hirsh.** 1988. *trans*-Spliced leader RNA exists as small nuclear ribonucleoprotein particles in *Caenorhabditis elegans*. Nature (London) **335**:556–559.
16. **Vournakis, J. N., A. Efstratiadis, and F. C. Kafatos.** 1975. Electrophoretic patterns of deadenylylated chorion and globin mRNAs. Proc. Natl. Acad. Sci. USA **72**:2959–2963.
17. **You, C.-Y., M. Ares, Jr., and A. M. Weiner.** 1985. Sequences required for 3′ end formation of human U_2 small nuclear RNA. Cell **42**:193–202.
18. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.