Vol. 211, No. 1, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
June 6, 1995 Pages 49-53

CIS-SPLICING AND POLYADENYLATION OF ACTIN RNA CAN PRECEDE 5' *TRANS*-SPLICING IN NEMATODES

Kuninori Shiwaku[1] and John E. Donelson*

Department of Biochemistry, University of Iowa and Howard Hughes Medical Institute, Iowa City, Iowa 52242

Received April 26, 1995

Summary: In nematodes many precursor RNAs undergo both *cis*-splicing (intron removal) and *trans*-splicing (spliced leader addition) via molecular mechanisms thought to be very similar. We examined the temporal order of *trans*- and *cis*-splicing by analyzing cDNAs of partially processed transcripts from an actin gene of the parasitic nematode, *Onchocerca volvulus*, which has five short introns. We found that in this model gene system *trans*-splicing can follow *cis*-splicing and polyadenylation, and we obtained no evidence for alternative pathways of splicing. This result is in contrast to the tubulin genes of African trypanosomes where *trans*-splicing precedes polyadenylation (1).

The primary transcripts of 70-80% of the protein-encoding genes in nematodes undergo both *trans*-splicing and *cis*-splicing during their subsequent maturation into mRNAs (2,3). During *trans*-splicing a 5' capped "spliced leader" (SL) of 22 nucleotides is transferred from the 5' end of its own precursor RNA, the SL RNA, to the 5' end of an RNA destined to become an mRNA. In *cis*-splicing the short introns of nematode genes (usually between 50 and 300 nucleotides (4)) are excised via mechanisms similar to those in other eukaryotes. Both splicing reactions are thought to occur through two trans-esterification steps that form a conventional 5'-3' phosphodiester bond between two exons and excise the intron sequences as either a lariat structure (*cis*-splicing) or a Y structure (*trans*-splicing), as reviewed in (3-6). The consensus 5' donor and 3' acceptor sites for both types of splicing conform to the GU/AG rule. Among the questions remaining about *trans*- and *cis*-splicing are: how do they co-exist in the same nucleus; how do the intermolecular *trans*-splicing substrates associate with each other; and what is the temporal order of the splicing reactions? Here we address the latter question in the case of the actin transcript of *Onchocerca volvulus*, where five introns must be removed and the 5' SL added during the formation of the actin mRNA (Fig. 1).

[1] Present address: Department of Environmental Medicine, Shimane Medical University, Shimane 693, Japan.

* Corresponding author. FAX: 319-335-6764.

0006-291X/95 $5.00

MATERIALS AND METHODS

Four different amplified cDNA libraries of poly(A)$^+$ RNA from adult *O. volvulus*, collectively containing a total of 34 million independent cDNA clones in bacteriophage lambda gt11 (7), were screened with two probes. Probe P1 was derived from the genomic DNA region immediately upstream of the SL addition site of one of the four actin genes in the diploid *O. volvulus* genome (8). The other probe was a nearly full length actin cDNA. Both probes were amplified by the polymerase chain reaction (PCR) and labeled with ^{32}P using a random DNA priming kit (Amersham Corp., Arlington Heights, IL). Libraries were plated at a density of about 20,000 phage per Petri dish (14 cm dia.). For primary screenings two replica filters of the plates were made, one to be hybridized with probe P1 and the other with actin cDNA. The inserts of phage that hybridized to both probes were subcloned into pBluescript plasmid and the subcloned inserts sequenced using T7 DNA polymerase (Sequenase, United States Biochemical Corp., Cleveland, Ohio). PCRs using primers containing the 22 nucleotide SL sequence and sequences complementary to the 3' ends of three introns of the actin gene were conducted for 25 cycles (9).

RESULTS AND DISCUSSION

O. volvulus, a parasitic nematode that causes onchocerciasis or river blindness in Africa and Latin America, is related phylogenetically to *Caenorhabditis elegans* and *Ascaris lumbricoides*, the two nematodes in which most of the studies on nematode *trans*-splicing have been conducted. While examining the organization and expression of the actin genes in *O. volvulus*, we showed that mature actin mRNAs contain a 5' SL at one of three potential acceptor sites located 7, 24 or 27 nucleotides in front of the start codon. We also identified an unexpected actin cDNA clone, called intermediate A1, that contains the complete actin coding sequence (376 codons), the 3' untranslated region (3' UTR) with a poly(A) tail, and an additional 514 nucleotides at its 5' end that are identical to the sequence of the genomic DNA region immediately upstream of the SL addition sites (9). Since the actin gene contains five introns, ranging from 122 to 207 bp, this cDNA appears to be derived from a partially processed actin precursor RNA in which the introns had been excised and polyadenylation had occurred, but the SL had yet to be added (Fig. 1).

To identify cDNAs of other partially processed actin transcripts, we screened about 1 million clones in four different *O. volvulus* cDNA libraries with probe P1 (Fig. 1) and an actin cDNA. About 0.2% of the clones hybridized to the actin cDNA probe and four of the cDNAs also hybridized to probe P1. These four cDNAs, named intermediates A2 - A5, were subcloned and analyzed. Each of these cDNAs were found to lack the SL and to have a long 5' region, the complete coding region and a 3' UTR (Fig. 1). Three of these four new cDNAs have a 3' poly(A) tail. Intermediate A3 has a 3' UTR that ends just upstream of the putative polyadenylation signal, AATAAA, but does not have a poly(A) tail, presumably because second strand cDNA synthesis was incomplete during the cDNA library construction.

Sequence determinations of intermediates A2 - A5 revealed that intermediate A2 contains a complete intron 4 but does not have the other introns, whereas intermediate A3 contains a complete intron 5 and does not have the other introns. Intermediates A4 and A5 resemble the original intermediate A1 in that none of the introns nor the SL is present, but a

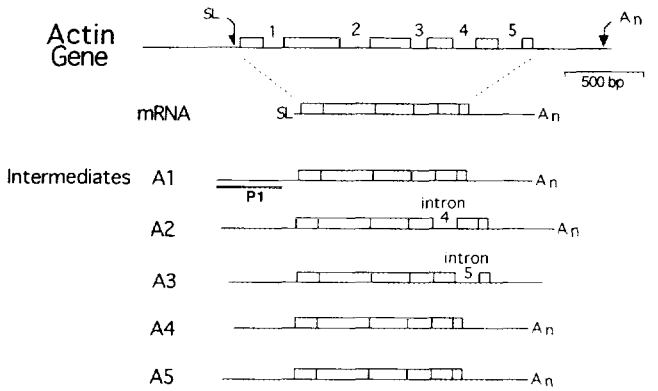

FIG. 1. Diagram showing the *O. volvulus* actin gene and mRNA, and five cDNAs, labeled Intermediates A1-A5, that are derived from partially processed actin precursor RNA molecules. The top line labeled Actin Gene depicts 3.1 kb of sequence given in ref. (9) and available in the GenBank database as accession no. M84916. The actin gene has six exons (rectangles) and five introns, labeled 1-5 (9). The sites of poly(A) addition and SL addition are shown. Hybridization probe P1 is indicated.

3' poly(A) occurs. Each of these five cDNAs are products of independent cloning events because they have different 5' boundaries with the vector and/or different numbers of A residues in the poly(A) tail.

As an alternative approach for finding cDNAs of partially processed actin transcripts, we conducted PCR amplifications using as the template total DNA extracted from aliquots of the cDNA libraries. The 5' primer for these PCR amplifications was the SL sequence, and the 3' primers were sequences complementary to the 3' ends of intron 3, intron 4 or intron 5. Primers derived from sequences in introns 1 and 2 were not used because they are too A/T-rich for the primers to bind specifically. The goal of these amplifications was to obtain evidence for actin pre-mRNAs that contained a 5' SL and still possessed at least one intron. However, in each of these different PCRs no amplification product was obtained, suggesting that the cDNA libraries do not have cDNA clones containing both the SL and either intron 3, 4 or 5 (data not shown). As a positive control for these experiments, PCR amplifications were conducted using a 5' primer derived from upstream of the SL addition site and 3' primers derived from introns 4 or 5. In each case a weak amplification product was observed (not shown), as expected since at least the intermediate clones A2 and A3, and perhaps others not detected during the screenings, would serve as a template for these amplifications. When the SL primer and a 3' primer from the actin coding region were used, a strong amplification product was obtained, reflecting the many full length actin cDNAs in the library.

Reverse transcriptase-PCRs with the same SL and intron primers were conducted using 5 μg of total RNA from adult *O. volvulus* as the template and reverse transcriptase as the enzyme in the first step of the amplification. Unfortunately, there is no animal model for *O. volvulus* infections so the parasites must be surgically excised from human volunteers in

Africa or Latin America, which severely limits the amount of parasite material available. As a result, we were unable to commit larger amounts of RNA for these reverse transcriptase PCRs and we did not obtain any new information from these experiments.

The following conclusions can be drawn from this limited amount of primary data. First, although the 5' end of the nascent precursor RNA potentially is available for *trans*-splicing during synthesis of the rest of the RNA, *cis*-splicing and 3' polyadenylation both can occur, and indeed can be completed, before *trans*-splicing occurs. Furthermore, the presence of the 3' poly(A) on four of the five cDNAs suggests that *trans*- and *cis*-splicing both can occur after release of the precursor RNA from the RNA polymerase II transcription complex (10). Although the RNA polymerase II complex contains high concentrations of *cis*-splicing factors, these could be simply storage sites for the splicing components (11).

The fact that one of the cDNAs contains intron 4 and another has intron 5 suggests that the *cis*-splicing apparatus is not constrained in a processive 5' - 3' manner to an obligatory pathway. It has been suggested that as particular introns are excised, the conformation of the RNA changes and new splicing sites become available (12). If such is the case here, the conformation of the *O. volvulus* actin pre-mRNA could be influenced by the pre-mRNA sequence and the *cis*-spliceosome without direct influence of the RNA polymerase II and the poly(A) addition complex.

One component that the nematode SL provides to the mRNAs is a 2,2,7-trimethyl-guanosine-cap structure that in most other eukaryotic systems is found only at the 5' ends of the snRNAs and not on the mRNAs (13). Since our data show that pre-mRNAs containing a poly(A) tail but lacking an SL can undergo intron excision, *cis*-splicing and polyadenylation in nematodes can occur without this cap structure at the 5' end of the pre-mRNA. This property contrasts with the requirement for the conventional mRNA cap structure, $m^7G(5')ppp(5')N$, in *cis*-splicing of mammalian cells (14). It may also be worth noting that our data do not address the fate of the pre-mRNA molecules that were the templates for synthesis of intermediate cDNAs A1 - A5. These pre-mRNAs might be incapable of undergoing further processing into mature mRNAs containing the 5' SL. However, arguing against this possibility is the fact that we were unable by PCR amplification to detect pre-mRNAs or intermediate cDNAs containing a 5' SL and at least one intron, suggesting that such intermediates are less abundant than the $poly(A)^+$ pre-mRNAs lacking both an SL and introns, whose cDNAs we did detect.

Why might *trans*-splicing follow *cis*-splicing in nematodes, despite the similarity in the molecular mechanisms responsible for both splicing events? The simplest explanation is that the substrates of *trans*-splicing are unlinked whereas those of *cis*-splicing are tethered. Thus, the intermolecular event of the donor and acceptor sites finding each other in *trans*-splicing likely takes longer than the corresponding intramolecular event in *cis*-splicing. Alternatively, another distinguishing feature between the two types of splicing is that the SL RNA serves as its own U1 snRNA in the *trans*-splicing reaction, whereas a conventional

U1 snRNA participates in *cis*-splicing (3,4). Perhaps this, or another, unique characteristic renders *trans*-splicing less efficient than *cis*-splicing. Still another consideration is that adjacent genes in nematodes can be transcribed into polycistronic precursor RNAs (2), and a downstream coding sequence might experience a different order of *trans*- and *cis*-splicing than the 5' coding sequence. In the haploid genome of *O. volvulus* two actin genes are separated by about 3 kb but it is not known if their sequences are part of the same precursor RNA.

Finally, the results described here are in direct contrast to the finding in African trypanosomes that *trans*-splicing precedes polyadenylation in the processing of tubulin pre-mRNAs (1). However, they are consistent with the observation in these same organisms that *trans*-splicing follows polyadenylation of HSP70 pre-mRNAs (15). Protein-encoding genes of African trypanosome genes do not have introns but they can be transcribed into polycistronic precursor RNAs as long as 60 kb (16), so perhaps these features also affect the relative order of *trans*-splicing and polyadenylation of different transcripts in these protozoa.

REFERENCES

1. Ullu, E., Matthews, K.R., and Tschudi, C. (1993) Mol. Cell. Biol. 13, 720-725.
2. Zorio, D.A.R., Cheng, N.N., Blumenthal, T., and Spieth, J. (1994) Nature 372, 270-272.
3. Nilsen, T.W. (1993) Ann. Rev. Microbiol. 47, 413-440.
4. Blumenthal, T. and Thomas, J. (1988) Trends in Genetics 4, 305-308.
5. Bonen, L. (1993) FASEB J. 7, 40-46.
6. Donelson, J.E. and Zeng, W. (1990) Parasitol. Today 6, 327-334.
7. Donelson, J.E., Duke, B.O.L., Moser, D., Zeng, W., Erondu, N.E., Lucius, R., Renz, A., Karam, M., and Flores, G.Z. (1988) Mol. Biochem. Parasitol. 31, 241-250.
8. Zeng, W., Alarcon, C.M., and Donelson, J.E. (1992) Mol. Cell Biol. 10, 2765-2773.
9. Zeng, W. and Donelson, J.E. (1992) Mol. Biochem. Parasitol. 55, 207-216.
10. Gunderson, S.I., Beyer, K., Keller, W., Boelens, W.C., and Mattaj, L.W. (1994) Cell 76, 531-541.
11. Rosbash, M. and Singer, R.H. (1993) Cell 75, 399-401.
12. Lewin, B. (1994) Genes V, p. 911. Oxford University Press, Oxford, England.
13. Maroney, P.A., Hannon, G.J., and Nilsen, T.W. (1990) Proc. Natl. Acad. Sci. USA 87, 709-713.
14. Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994) Cell 78, 657-668.
15. Huang, J. and Van der Ploeg, L.H.T. (1991) Mol. Cell. Biol. 11, 3180-3190.
16. Johnson, P.J., Kooter, J.M., and Borst, P. (1987) Cell 51, 273-281.