

## Gene linkage and steady state RNAs suggest *trans*-splicing may be associated with a polycistronic transcript in *Schistosoma mansoni*<sup>1</sup>

Richard E. Davis \*, Scott Hodgson

Department of Biological Sciences, Fordham University, Bronx NY, 10458, USA

Received 10 March 1997; accepted 24 June 1997

### Abstract

Spliced leader (SL) *trans*-splicing generates the 5' end of mature mRNAs through the addition of a small exon to pre-mRNAs in some flagellates (kinetoplastida and euglenoids) and metazoans (nematodes and flatworms). Although SL addition in the kinetoplastida and a subset of nematode genes serves to resolve multicistronic mRNAs into monocistronic, capped mRNAs, information regarding the functional significance of *trans*-splicing in flatworms is limited. We describe here the identification and characterization of a closely linked gene upstream from the *trans*-spliced enolase gene in the flatworm *Schistosoma mansoni*. This gene produces a non-*trans*-spliced mRNA encoding a ubiquinol binding protein, UbCRBP, that is a component of the ubiquinol-cytochrome C reductase complex. The distance between the UbCRBP polyadenylation site and the enolase *trans*-splice acceptor site is exceptionally short, only 54 nucleotides. Primer extension (5' RACE), RT-PCR, and RNase mapping have identified steady state, *cis*-spliced RNAs which significantly overlap both the UbCRBP and enolase genes. These transcripts contain the 5' ends of mature UbCRBP mRNAs, extend through UbCRBP, across the intergenic region, and a significant distance 3' into the enolase gene. Interestingly, the close linkage between the UbCRBP and enolase genes is conserved in a second flatworm, *Fasciola hepatica*, which also *trans*-splices the downstream enolase gene. Taken together, the role of SL addition in resolving multicistronic transcripts in both *C. elegans* and the kinetoplastida, the conservation of UbCRBP/enolase gene linkage in two divergent trematodes, and the multicistronic organization of schistosome UbCRBP/enolase RNAs are consistent with the suggestion that these two genes are likely to be cotranscribed and that *trans*-splicing in flatworms may be associated with polycistronic transcripts. © 1997 Elsevier Science B.V.

**Abbreviations:** SL, spliced leader; SL RNA, spliced leader RNA; ORF, open reading frame; UbCRBP, ubiquinol cytochrome reductase complex binding protein.

\* Corresponding author. Tel.: +1 718 8173630; Fax: +1 718 8173645; E-mail: rdavis@murray.fordham.edu

<sup>1</sup> **Note:** Nucleotide sequence data reported in this paper are available in the EMBL, GenBank, and DDJB data bases under the accession numbers: U90319 and U91748.

**Keywords:** *Trans*-splicing; Pre-mRNA splicing; Spliced leader; *Trans*-splice acceptor; Trematodes; Flatworms; Enolase; UbCRBP; Polycistronic transcript

## 1. Introduction

*Trans*-splicing accurately joins exons derived from independently transcribed RNAs. In spliced leader *trans*-splicing, a leader sequence (the spliced leader, SL) is donated from the 5' end of a small, non-polyadenylated RNA, the spliced leader RNA (SL RNA), to pre-mRNAs to form the 5' terminal exon of mature mRNAs (see recent reviews [1–3] and references therein). SL addition occurs in flagellate protozoa (kinetoplastida and euglenoids) and in two groups of early metazoa, the nematodes and flatworms [4–8]. The presence of *trans*-splicing in both roundworms and flatworms suggests that this form of RNA processing might be an evolutionarily important and common form of gene expression in early metazoa. The overall distribution of *trans*-splicing, its evolutionary origin, and its functional significance in several organisms, however, remains unknown. SL addition is of interest in flatworms (Phylum Platyhelminthes) as these early metazoa are likely to have been the first bilateral animals, and in some evolutionary trees, a flatworm-like ancestor occupies an important phylogenetic position as the putative progenitor of a number of early invertebrate groups [9–12].

All mRNAs in kinetoplastid flagellates are thought to acquire a SL during maturation. Transcription units in the kinetoplastida are typically polycistronic and *trans*-splicing in these organisms serves at least two functions: (1) providing a cap for the mRNA and (2) resolving polycistronic primary transcripts into monocistronic mRNAs (see [3] and references therein).

In the nematodes *Caenorhabditis elegans* and *Ascaris suum*, approximately 70% and 80–90%, respectively, of their mRNAs are generated by *trans*-splicing [13,14]. Although considerable attention has been paid to the mechanism and determinants of *trans*-splicing in nematodes, the biological significance of the majority of nematode *trans*-splicing (SL1) remains unknown

[1,2,15]. However, a subset of *trans*-spliced genes (~15% of all mRNAs) in *C. elegans* acquire a second type of spliced leader, SL2 [16], which contributes toward resolving polycistronic transcripts into individual mRNAs [1,13,17]. The overall frequency and distribution of polycistronic transcripts and specialized SLs such as SL2 in nematodes other than *C. elegans* is currently not known.

In contrast with nematodes, only a relatively small proportion of mRNAs are thought to be *trans*-spliced in the flatworm *Schistosoma mansoni* [18]. Previous analyses in *S. mansoni* led us to conclude that the functional significance of *trans*-splicing in flatworms is most likely associated either with some property conferred on recipient mRNAs by addition of the SL or related to some aspect of the transcription and/or primary transcripts of *trans*-spliced genes [18]. During the course of characterizing further the *trans*-spliced enolase gene in *S. mansoni*, we have identified a closely linked upstream open reading frame encoded on the same strand as the enolase mRNA. We show here that non-*trans*-spliced mRNAs containing this ORF are expressed in adult schistosomes. These mRNAs encode a component of the ubiquinol-cytochrome C reductase complex, the so-called ubiquinol binding protein (UbCRBP). The site of UbCRBP polyadenylation is only 54 nucleotides upstream of the enolase *trans*-splice acceptor site. This gene linkage and *trans*-splicing of the downstream enolase gene is conserved in a second highly diverged trematode, *Fasciola hepatica*. Overall, our data including the organization of steady state RNAs derived from the locus and the gene linkage, combined with the precedence for the role of *trans*-splicing in resolving multicistronic transcripts in the kinetoplastida and nematodes, is consistent with cotranscription of UbCRBP and enolase in schistosomes. The data therefore suggest that *trans*-splicing in flatworms may serve to resolve polycistronic transcripts into monocistronic mRNAs.

## 2. Materials and methods

### 2.1. Organisms

Frozen adult *Schistosoma mansoni* and *Fasciola hepatica* were kindly provided by Ron Blanton (Case Western Reserve University) and Allison Rice-Ficht (Texas A&M University), respectively.

### 2.2. Nucleic acid isolation and blot analyses

Genomic DNA was isolated from frozen worms powdered on dry ice as described [19]. Total RNA was purified either by guanidinium-hot phenol [19] or one of several variations of acid-phenol extraction [20,21] (TRIzol Reagent, GIBCO BRL/Life, Gaithersburg, MD). Poly (A) + mRNAs were selected by oligo(dT)-cellulose [19] or biotinylated oligo(dT) and streptavidin-coated paramagnetic particles (PolyATtract mRNA Isolation, Promega, Madison, WI). Northern and Southern blots, probe preparation, and hybridization and washing conditions were as described [7,19].

### 2.3. Primer extension analysis and rapid amplification of cDNA ends

Primer extension was performed as described [19] using end-labelled Hpa II pBR322 and DNA sequencing reactions as molecular size markers. The 5' rapid amplification of cDNA ends (RACE) was performed with the 5' RACE System from Gibco BRL (Gaithersburg, MD; according to the manufacturers instructions) or using an alternative protocol involving removal of the mRNA cap and ligation of an RNA oligonucleotide to the 5' end of the mRNA prior to first strand cDNA synthesis as described [22–25]. In brief, total RNA was dephosphorylated (Boehringer Mannheim, Indianapolis, IN), decapped using tobacco acid pyrophosphatase (Epicentre, Madison, WI), and an RNA oligonucleotide (AUAGGACUAACUACACUGAGGAUCCAAA) ligated to the decapped and dephosphorylated RNAs with T4 RNA ligase (New England BioLabs, Beverly, MA) using conditions described by the manufacturers. To obtain the 3' end of the UBCRB mRNA, 3' RACE was performed as described [26]

with the modification of Rother [27] as previously described [18]. The PCR products were gel purified using QIAEX gel extraction (Qiagen, Chatsworth, CA) and directionally cloned using ligation-independent cloning with either the PCR-Direct (Clontech, Palo Alto, CA) or the pAMP cloning system (Life Technologies) as described by the manufacturers. Blunt PCR products were also cloned into Bluescribe/Bluescript plasmid vectors (Stratagene Cloning Systems, La Jolla, CA).

### 2.4. 5' flanking anchor PCR of the *Fasciola hepatica* enolase gene

Genomic DNA was partially digested with Eco RV and sequence extending 2.7 kb upstream from *Fasciola* enolase trans-splice acceptor site was amplified and cloned using the PromoterFinder Construction Kit as described by the manufacturer (Clontech).

### 2.5. RT-PCR and RNase mapping

Genomic DNA was removed from RNA samples by treatment with 1 unit RQ-1 DNase I (Promega, Madison, WI) per 4 µg of total or 0.5 µg of poly(A) + RNA for 30 min at 37°C followed by phenol/chloroform extraction and ethanol precipitation prior to reverse transcription or RNase mapping. cDNA was synthesized from total and poly(A) + RNA using random primers, gene specific primers, or oligo(dT) with superscript II reverse transcriptase (Gibco BRL) using the conditions recommended by the manufacturer. PCR amplified products were RNA and reverse transcriptase dependent. RNase mapping with the cloned PCR products was carried out using RNAase A and T1 with 2–50 µg of total RNA or 1–5 µg of poly A + RNA essentially as described [28] or using the HybSpeed RPA kit as described by the manufacturer (Ambion, Austin, TX).

### 2.6. Sequencing and sequence analysis

Plasmid DNA was prepared as previously described [7,18,19] or using QIAprep (QIAGEN, Chatsworth, CA). Clones were sequenced by the



Fig. 1.

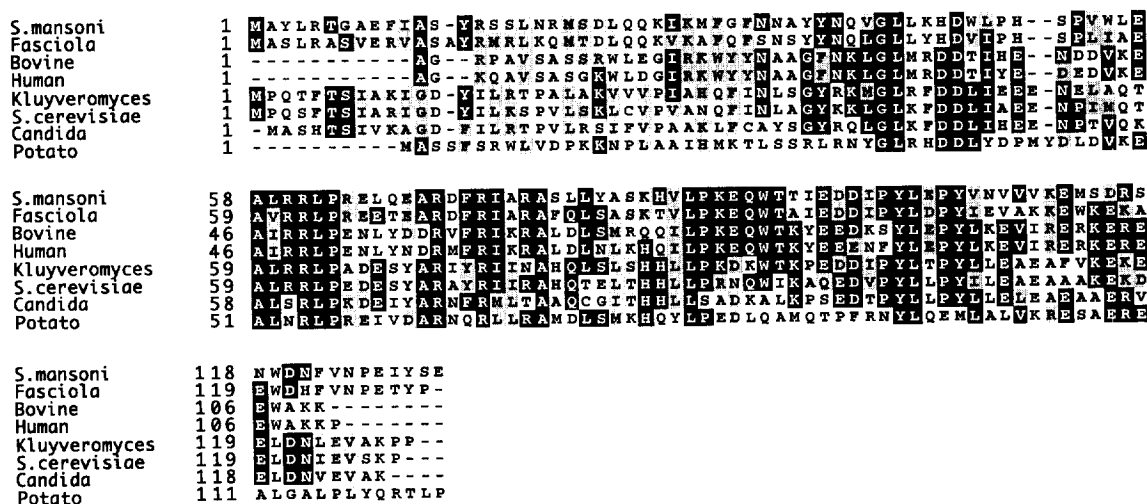


Fig. 1. Closely linked gene upstream of enolase encodes UbCRBP. (A) *Schistosoma mansoni* UbCRBP gene sequence extending into exon 3 of the downstream enolase gene (exon 1 of the enolase gene is the SL). Lower case letters represent upstream regions, introns, or the intergenic region between UbCRBP and enolase. Uppercase letters represent exon sequences (light shading, UbCRBP; darker shading, enolase) and the predicted translation products of the encoded mRNAs are shown below the exon sequence. The putative polyadenylation signal, intergenic region, and *trans*-splice acceptor site for enolase are marked. The alternate splice donor sites contributing to the definition of UbCRBP exon 1 are shown with triangles with the GT dinucleotide underlined. Polyadenylation sites determined by 3' RACE are marked with diamonds and the number of clones identified corresponding to each site are provided in parentheses. (B) Linkage between UbCRBP and the *trans*-spliced enolase gene in *Fasciola hepatica*. Only the 3' end of *Fasciola* UbCRBP (accession No. U91748) and a portion of enolase exon 2 are shown (exon 1 is the *Fasciola* spliced leader [37]). Letters and shading are the same as in (A). (C) Multiple alignment of the upstream schistosome and *Fasciola* ORFs with phylogenetically diverse UbCRBP proteins. Dashes indicate gaps used to generate the optimal alignment, dark shading indicates identical residues, and light shading indicates conservative amino acid substitutions.

dideoxynucleotide method on alkali-denatured plasmids using the USB Sequenase Kit (USB, Cleveland, OH) as described [7,19] or by cycle sequencing (AmpliCycle: Perkin Elmer, Foster City, CA;  $\Delta$ -Genozyme System: Genomymx, Foster City, CA; Thermosequenase: USB) as described by the manufacturer.

Sequence analysis and primer design were performed with MacVector (Kodak, Rochester, NY), GeneWorks (Intelligenetics, Mountain View, CA), Genetics Computer Group (GCG) (Madison, WI), Oligo 4.0 (NBI, Plymouth, MN), and BLAST programs [29].

### 3. Results

### 3.1. Closely linked gene upstream of enolase encodes UbCRBP

We previously identified several *trans*-spliced genes in *Schistosoma mansoni* including one which encodes a protein with significant similarity to known enolases [18]. Characterization of sequences upstream of enolase exons (Fig. 1A) identified an open reading frame (ORF) that terminates only 117 nucleotides 5' of the *trans*-splice acceptor site for the downstream enolase gene

[18]. Comparison of the conceptual translation of this ORF with protein databases indicated that the 130 amino acid protein has significant similarity to a component of the ubiquinol-cytochrome C reductase complex, the ubiquinol binding protein, UbCRBP (Fig. 1C). Amplification and sequence analysis of 5' flanking sequences for the *trans*-spliced enolase gene in another flatworm indicated that UbCRBP is also closely linked with enolase in *Fasciola hepatica* (Fig. 1B and C). The *Fasciola* UbCRBP ORF terminates 110 nucleotides upstream of the enolase *trans*-splice acceptor site (Fig. 1B).

In order to determine if an mRNA encoding the UbCRBP ORF was expressed from this locus in adult schistosomes, we carried out Northern blot analysis (Fig. 2A). A probe corresponding to the UbCRBP ORF hybridized to a relatively abundant mRNA of ~550 bases in total RNA (Fig. 2A, lane 2). The steady state levels of the UbCRBP and enolase mRNAs in adult schistosomes are roughly similar as shown by the hybridization in Fig. 2A, lane 3. Probes extending either 4 kb further upstream of the UbCRBP locus (Fig. 2A, lane 1) or 1.5 kb downstream (Fig. 2A, lane 4) from the site of polyadenylation of the enolase mRNA did not hybridize to other mRNAs of similar abundance, although a small RNA (the gene encoding this RNA does not hybridize with an SL RNA probe) is encoded in the upstream region (Fig. 2).

Southern blot analyses (data not shown), extensive characterization of genomic clones corresponding to these genes [18], and PCR and RNase mapping studies (see below) indicate that the UbCRBP and enolase genes are encoded by a single locus in the schistosome genome.

### 3.2. Characterization of UbCRBP mRNAs

To further define and characterize the mature UbCRBP mRNA, we performed 5' and 3' RACE on total RNA. The 3' RACE analyses demonstrated that the entire predicted coding

region matched the gene sequence and that polyadenylation can occur at several positions within a region of ~8 nucleotides (Fig. 1A; nucleotides 1941–1947). The exact location of cleavage and polyadenylation cannot be identified in several cases due to polyadenylation occurring at an A or within a run of A residues in the genomic sequence. Nevertheless, the dominant polyadenylation site (60% of mRNAs characterized) is only 54 nucleotides upstream from the enolase *trans*-splice acceptor site.

Comparison of 5' RACE products (Fig. 3(A)) with genomic sequence (Fig. 1A) indicates that UbCRBP mRNAs are not *trans*-spliced. Furthermore, UbCRBP mRNAs are encoded by two exons, the first of which is non-coding and utilizes at least four alternative splice donor sites. While we recognize the potential limitations, we assume here that the 5' ends of the RACE products represent the mature 5' ends of UbCRBP mRNAs and that the variation at the 5' ends of these RACE products is most likely representative of heterogeneous transcription initiation. Similar heterogeneity and variation at the 5' end of UbCRBP RNAs was also observed in both primer extension (data not shown) and RNase mapping experiments (see below). Furthermore, the relative spacing between the 5' end of individual clones and the actual exon 1 splice donor site selected (Fig. 3A) fits a pattern suggestive of a minimum exon size requirement for splicing. This pattern also tends to support the likelihood of heterogeneous transcription initiation. All of the mature UbCRBP mRNAs contain the same second exon (except for the exact site of polyadenylation) and code for the same conceptual translation product.

### 3.3. RNase mapping the mature 5' end of a UbCRBP mRNA

We used RNase protection (Fig. 4(A)) to independently identify the 5' ends of mature UbCRBP mRNAs corresponding to clones N3, N4, N8, N9, and P1 which use the 5' most al-

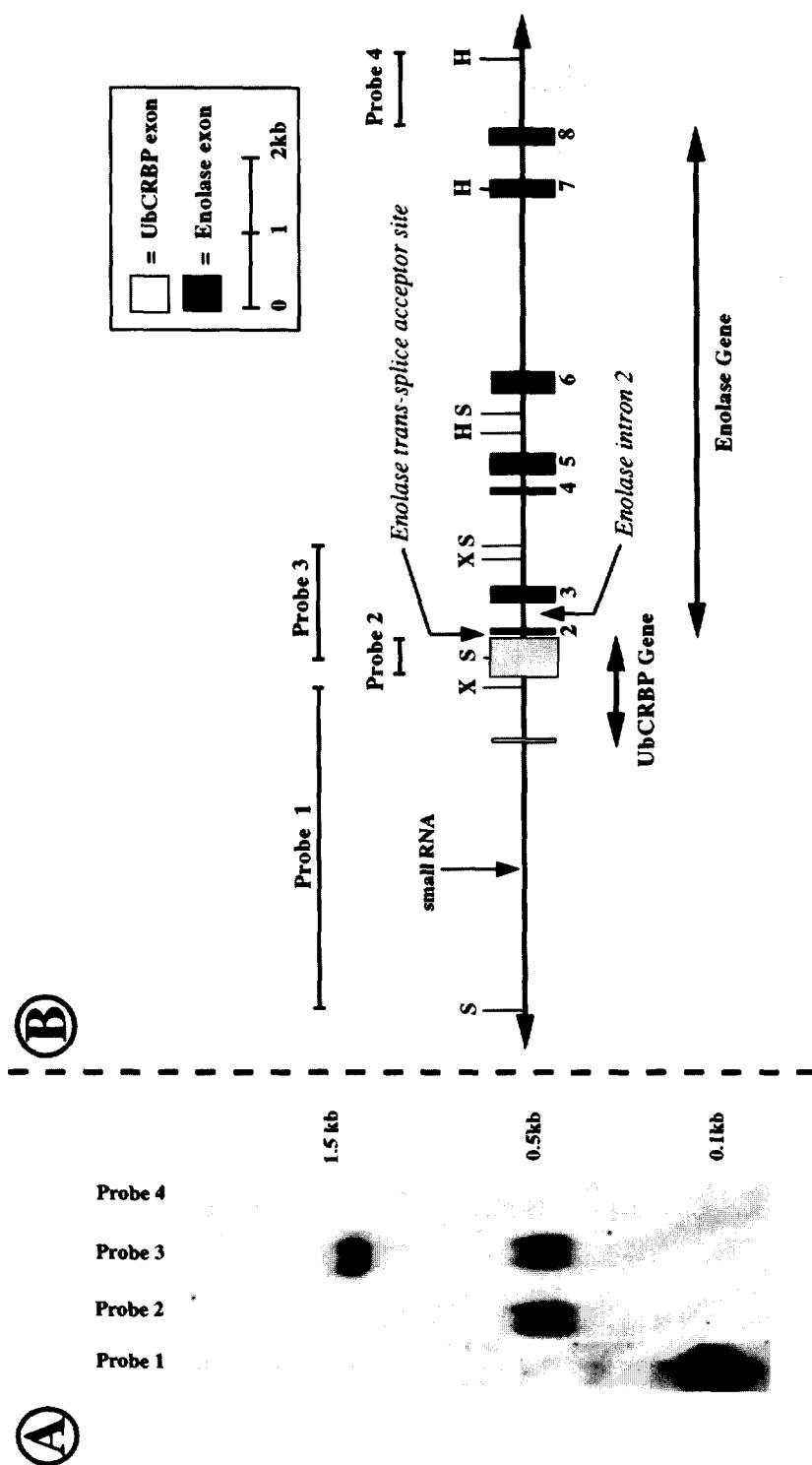


Fig. 2. UbCRBP ORF is expressed in a 550 base mRNA. (A) Agarose Northern blots of total RNA were hybridized with random primer labelled probes (see probes in 2B) corresponding to sequences extending 4000 bp upstream from UbCRBP exon 2 (Probe 1, Sal-Xba fragment), the protein coding region of UbCRBP (Probe 2), a 1.5 kb Sal I fragment consisting of portion of exon 2 of UbCRBP extending 3' through enolase exon 3 and into intron 3 of enolase (Probe 3), and an ~1.5 kb region 3' of the polyadenylation site of enolase (Probe 4). The hybridization in 1 is to a small RNA located as illustrated in 2B. The gene encoding this RNA does not hybridize to an SL RNA probe. (B) Schematic of the locus illustrating Northern blot probes and UbCRBP and enolase exons. For the enolase gene, exon 1 is the spliced leader exon and upstream of the *trans*-splice acceptor site is considered intron 1. H, HindIII; S, SalI; X, XbaI.

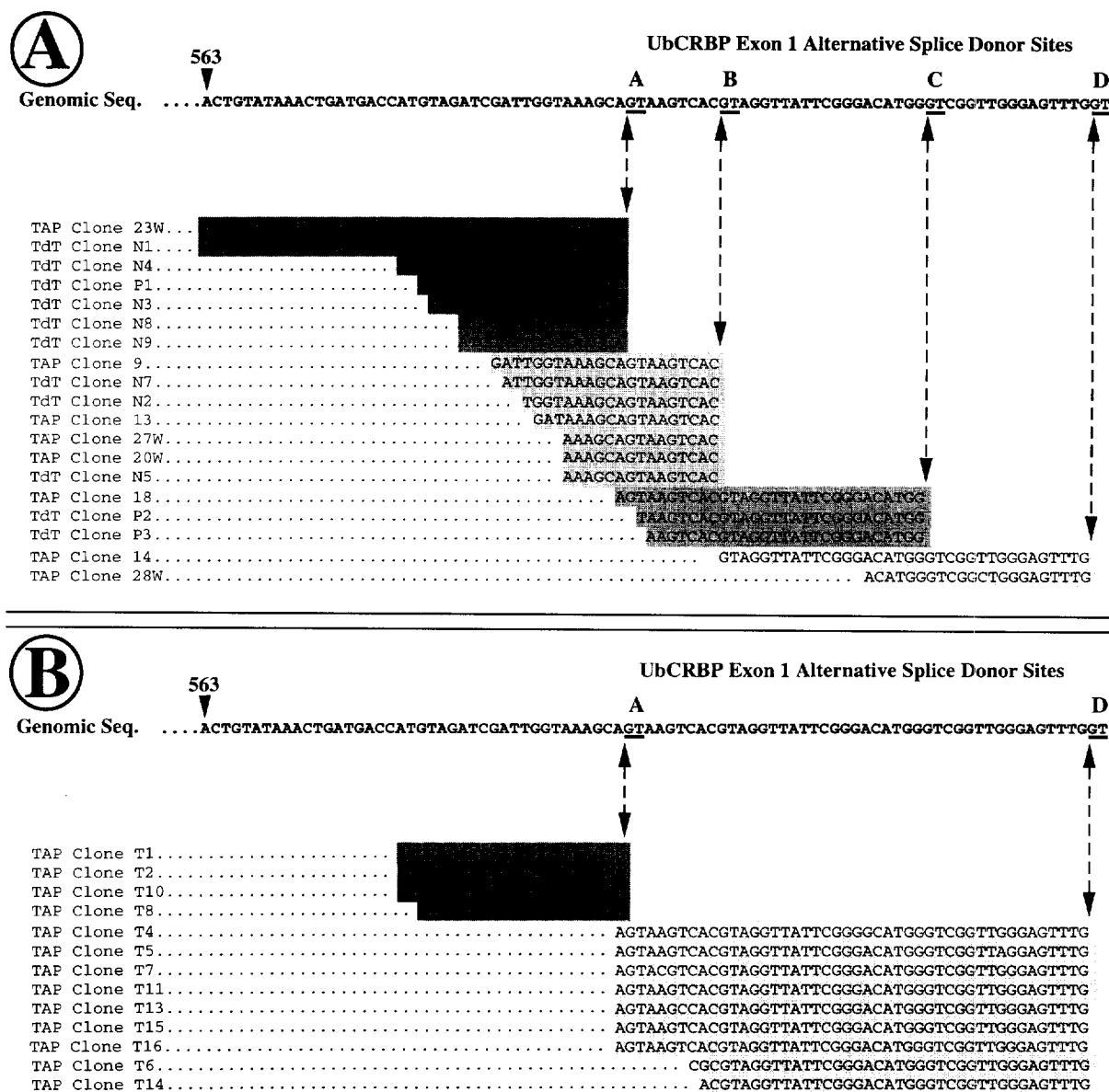


Fig. 3. The 5' ends and organization of exon 1 of UbCRBP RNAs obtained from 5' RACE. Sequences represent individual clones derived from two different 5' RACE methods as described in Section 2. TdT represents clones obtained by tailing first strand cDNA, and TAP represents clones obtained using decapping and ligation of an RNA oligonucleotides to the 5' end of the mRNA prior to first strand cDNA synthesis. (A)–(D) represent alternative splice donor sites used for splicing to a constant second UbCRBP exon. The corresponding nucleotide position (563) in the genomic sequence in Fig. 1 is provided. (A) UbCRBP 5' RACE clones. The clones were obtained using a 3' primer within exon 2 of UbCRBP (nucleotides 1839–1819) and consist of UbCRBP exon 1 and part of exon 2 extending 3' to the primer used in the PCR. (B) UbCRBP/enolase 5' RACE clones. The clones were obtained using a 3' primer within exon 3 of enolase. The organization of all of these clones consists of UbCRBP exon 1 spliced to exon 2, the 54 nucleotide intergenic region between UbCRBP and enolase, and enolase exon 2 spliced to exon 3 (Fig. 4B).



ternative exon 1 splice donor site (Fig. 3A). The protected products observed in this experiment were heterogeneous with 5' ends mapping between nucleotides 582 and 587. Minor protected products with 5' ends mapping between nucleotides 563 and 568 were also observed and are consistent with 5' RACE clones N1 and 23W (Fig. 3A). The consistency between these protected products and the UbCRBP 5' RACE clones suggests that transcription initiation may be heterogeneous for this gene.

#### 3.4. 5' RACE of non-trans-spliced enolase transcripts

To attempt to identify the 5' ends of non-trans-spliced enolase transcripts, we carried out 5' RACE analyses using specifically primed cDNA (total RNA primed with nucleotides 2569–2548 of enolase exon 3 in Fig. 1A and Fig. 2B; note that exon 1 is the SL). The cDNA was amplified enriching for long PCR products with a nested primer located further 5' in enolase exon 3 (nucleotides 2499–2477). The primary ethidium bromide stained PCR product corresponded in size (~210 nucleotides, Fig. 4B, lane 7) and sequence (data not shown) to the mature full-length enolase mRNA with a 5' terminal SL [7]. Larger and less abundant ethidium bromide stained products were also observed of ~800 and 1100 bp. Southern blot hybridization indicated these products, as well as less abundant RACE products of ~1600 and 1900 bp, contained the intergenic region between UbCRBP and enolase in addition to UbCRBP exons 1 and 2 (Fig. 4B). Sequence analysis of the ~800 bp band indicates it corresponds to an RNA with the following organization from the 5' end (Fig. 4B): (1) UbCRBP exon 1 spliced to exon 2, (2) the coding region of UbCRBP, (3) the intergenic region between UbCRBP and enolase, and (4) enolase exons 2 and 3 spliced together. The sequences of the 5' ends of the ~800 bp products are shown in Fig. 3B. Mature UbCRBP mRNAs with the two different first exons illustrated in Fig. 3B were also detected in RNase mapping experiments (Fig. 5) and in Northern blots using specific oligonucleotide probes corresponding to each of

the two exons (data not shown). Sequence analysis of the ~1100 bp product indicates that it is identical to the 800 bp product except the enolase intron is present (Fig. 4B). Although the larger 5' RACE products were not sequenced, based on their size and multiple hybridization analyses, they are consistent with a putative primary transcript from the locus (~1900 bp) and an RNA with the enolase intron removed, but the UbCRBP intron retained (~1600 bp) (Fig. 4B). 5' RACE products, consistent in size and hybridization with an enolase transcription initiation site downstream of UbCRBP exon 1, were not observed using several different probes to the intergenic region, UbCRBP exon 2, and the UbCRBP intron (Fig. 4B). RNase mapping (Fig. 5) also did not protect RNAs consistent with enolase transcription initiation within UbCRBP exon 2. In addition, 5' RACE and PCR experiments did not detect UbCRBP mRNAs with either UbCRBP intron sequences or a SL at their 5' ends, further suggesting that enolase transcription initiation within the UbCRBP intron is unlikely. RNase mapping for steady state RNAs greater than 1000 bases (e.g. primary transcripts or intron containing transcripts extending from UbCRBP exon 1 into exon 3 of enolase) was not feasible since it exceeded the technical limits of the assay.

#### 3.5. RNase mapping demonstrates RNAs spanning the intergenic region between UbCRBP and enolase

The 5' RACE data suggest enolase transcripts may originate at the 5' end of UbCRBP and extend into the enolase gene. To independently identify these RNAs, we carried out RNase mapping on total RNA using two of the 800 bp enolase 5' RACE clones described above. These clones, T14 and T8 (Fig. 3B), represent transcripts with two variations in UbCRBP exon 1 and are otherwise identical extending 3' across the intergenic region into enolase exons (see schematic in Fig. 5A and B). RNAs with an organization identical to these clones (see products A1 and B1 in Fig. 5) and RNAs differing only in UbCRBP exon 1 (see products A2 and B2 in

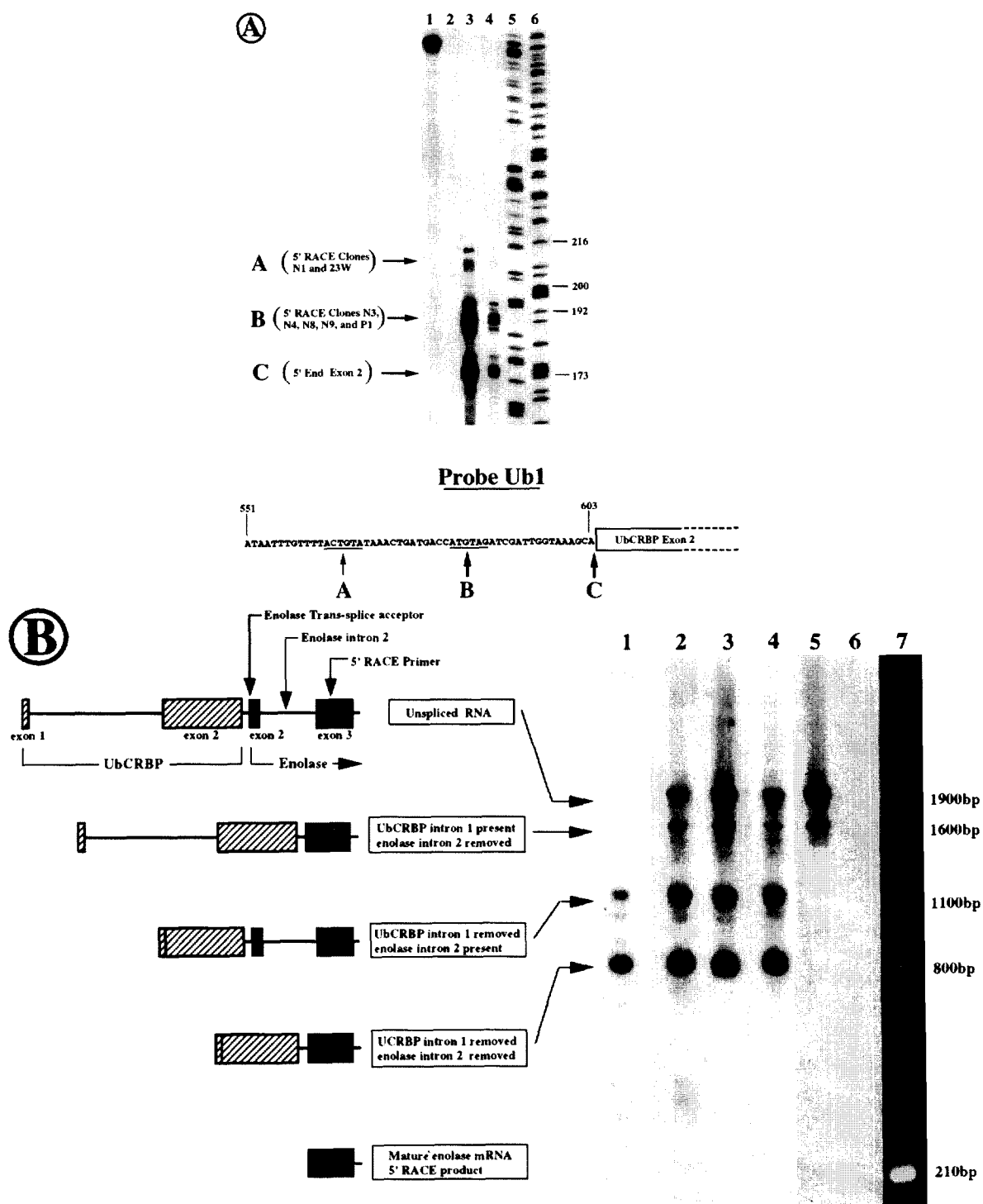


Fig. 4.

Fig. 5) were protected in these experiments. In addition, protected products corresponding to mature UbCRBP mRNAs (see products A3 and B3 in Fig. 5) were also observed, including the UbCRBP mRNAs which differ in the first exon of UbCRBP and use alternate 5' splice sites (see shorter products A4 and B4 in Fig. 5). Both the RACE and RNase mapping products were RNA dependent and cannot be detected in DNA.

### 3.6. *Trans-spliced enolase RNAs containing 5' introns are not detected*

RT-PCR experiments (data not shown) using an SL primer in combination with primers either in enolase exons 3, 4 or 5 (the SL is considered enolase exon 1; see enolase exons in Fig. 2B) readily amplified the mature enolase mRNA, but did not detect *trans*-spliced RNAs retaining enolase introns 2, 3 or 4. In contrast, a UbCRBP exon 2 primer in combination with these enolase primers detected RNAs consistent with putative primary transcripts and *cis*-splicing intermediates as well as RNAs with complete removal of enolase introns 2–4. The identity of these RT-PCR products were confirmed by cloning and sequencing. These products were then used in RNase mapping experiments (data not shown) with total and poly(A<sup>+</sup>) mRNA. Protected RNAs consistent in organization with UbCRBP exon 2 joined to the intergenic region and enolase exons 2–4

with the intervening enolase introns removed were readily detected in both total and polyadenylated RNA. These data, RNase mapping using as a probe UbCRBP exon 1 extending 3' into enolase exon 3 (described above), and the 5' RACE experiments further suggest that transcripts which contain UbCRBP extend a significant distance 3' into the enolase gene.

## 4. Discussion

A gene encoding ubiquinol-cytochrome C reductase binding protein, UbCRBP, is present on the same coding strand and located immediately upstream from the *trans*-spliced enolase gene in *Schistosoma mansoni*. Both genes are expressed in adult schistosomes and similar levels of the two steady state mRNAs accumulate. The UbCRBP gene produces several mRNAs which vary primarily in the organization of their 5' ends. The different 5' ends of these mRNAs likely result from a combination of heterogeneity in transcription initiation and several alternative exon 1 splice donor sites. The primary site of polyadenylation for the UbCRBP mRNAs is located only 54 nucleotides upstream from the *trans*-splice acceptor site for enolase.

Using a variety of techniques we have detected steady state RNAs that contain both UbCRBP and enolase sequences. These transcripts begin with sequences corresponding to the 5' ends of

Fig. 4. The 5' ends of mature UbCRBP mRNAs and non-*trans*-spliced enolase transcripts map to exon 1 of UbCRBP. (A) RNase mapping the 5' end of a mature UbCRBP mRNA. DNase I treated total RNA was protected with a probe consisting of a portion of exon 2 (to nucleotide 1641 in Fig. 1) extending upstream through exon 1 (exon 1 using the first 5' splice donor site in the gene; clone N4 in Fig. 3A) and into 5' flanking genomic sequence to nucleotide 551. The apparent site of transcription initiation is slightly heterogeneous as observed in the RACE products and corresponds to the clones shown (Fig. 3A). The shorter protected RNA fragment (C) corresponds to UbCRBP mRNAs using 3' splice donor sites for exon 1. Consequently, they map to the 5' boundary of exon 2. Lanes: 1, undigested probe; 2, probe digested with tRNA; 3, total RNA (4-fold greater exposure than lane 4); 4, total RNA; 5 and 6, ddA and ddT sequencing ladders, respectively. (B) RACE to identify the 5' ends of non-*trans*-spliced enolase transcripts. PCR products were resolved on agarose gels with ethidium bromide (lane 7), Southern blotted, and hybridized with either a probe corresponding to the UbCRBP/enolase intergenic region (nucleotides 1948–1966) (lanes 1–2; lane 1 is a lighter exposure), the 3' end of UbCRBP exon 1 (nucleotides 629–646) (lane 3), UbCRBP exon 2 (nucleotides 1926–1905) (lane 4), or the 3' end of the UbCRBP intron (nucleotides 1453–1473) (lane 5). Lane 6 illustrates representative hybridization following amplification with primers alone. The mature enolase 5' RACE product is 210 bp. The region upstream of the enolase *trans*-splice acceptor site is considered enolase intron 1.

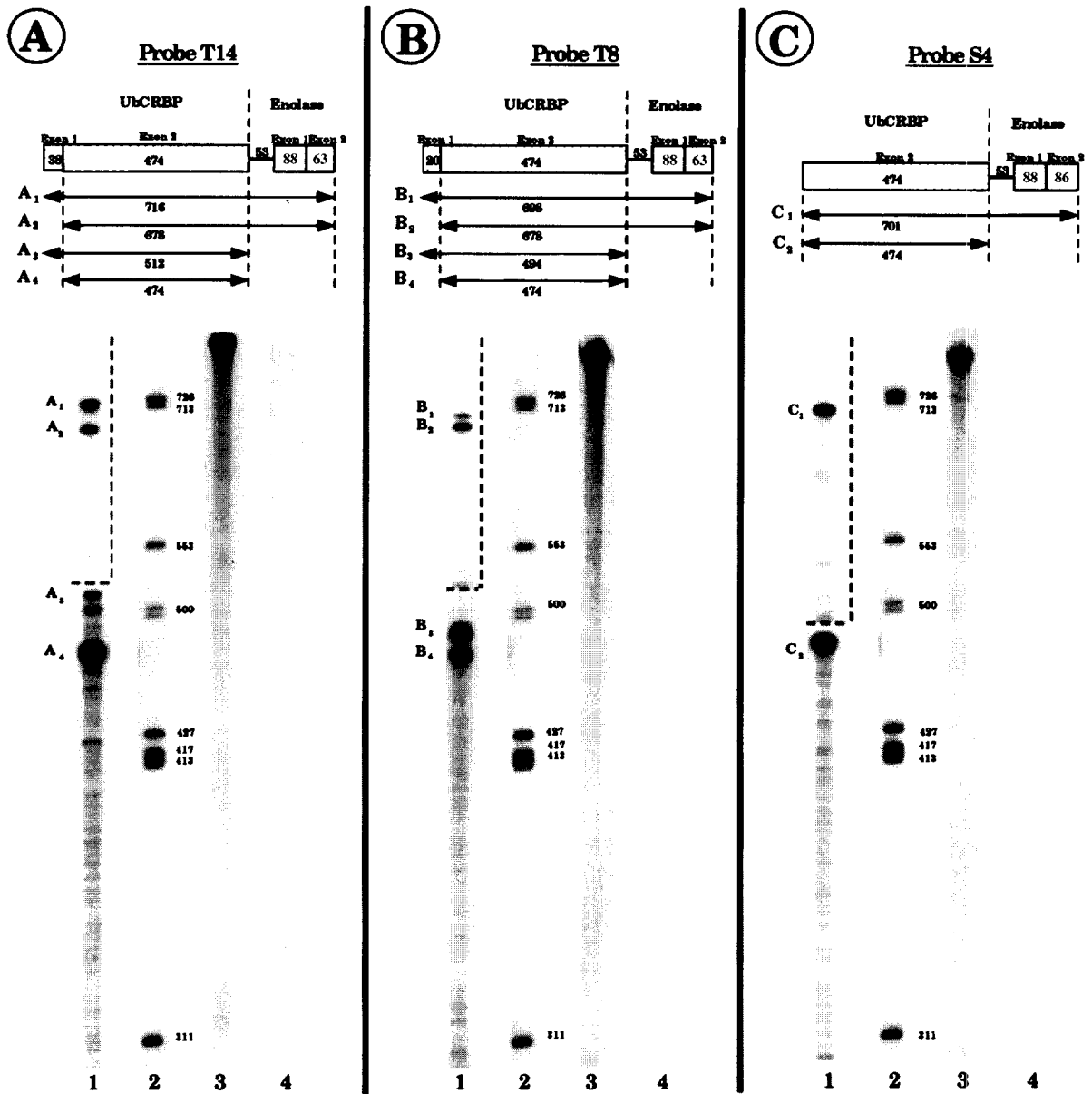


Fig. 5. RNase mapping demonstrates RNA transcripts spanning the two genes. DNase I treated total RNA was protected with antisense RNA probes corresponding to the diagrams in (A)–(C). Plasmid sequence was present extending both 5' and 3' of the probes shown. Probes in (A) and (B) were derived respectively, from clones T14 and T8 in Fig. 3B. The S4 probe in (C) served as an internal control and size marker in these experiments protecting all RNAs containing UbCRBP exon 2. The S4 was generated from a cloned RT-PCR product beginning at the first base of UbCRBP exon 2 and extending 3' across the intergenic region through enolase exon 2 into exon 3. Lanes: 1, protected *S. mansoni* RNAs; 2, end-labelled fX174–Hinf I markers; 3, undigested probe; and 4, probe digested with tRNA. The upper portion of lane 1 (marked by dashed lines) in each section represents a 4-fold longer exposure than the corresponding lower section.

mature UbCRBP mRNAs and extend through the intergenic region 3' at least 2000 nucleotides into the enolase gene. The major transcripts observed represent RNAs in which both the UbCRBP and enolase introns have been removed. However, RNAs with either both introns present or only one intron removed were also detected in hybridization experiments.

We are well aware of the limitations of solely analyzing steady state RNAs, particularly in the context of describing a precursor/product relationship or a productive pathway of mRNA generation. Unfortunately, additional methods including transfection and run-on transcription are currently not available in these human parasites. One possible interpretation of the RNase mapping and 5' RACE data is that initiation of both enolase and UbCRBP transcripts originates at the 5' end of UbCRBP exon 1. Alternatively, for these transcripts which extend 3' into the enolase gene, it is also possible that the sites which map to the 5' end of UbCRBP exon 1 represent hypersensitive sites in precursor RNA cleavage/processing events upstream of the UbCRBP gene. Nevertheless, the current data are consistent with transcripts derived from this locus which contain sequences corresponding to the 5' ends of mature UbCRBP mRNAs which extend 3' through the intergenic region at least 2000 nucleotides into the enolase gene. This results either in a bicistronic (or multicistronic) transcript or completely overlapping UbCRBP and enolase transcription units. A variety of data in the kinetoplastida and *C. elegans* (SL2 *trans*-spliced genes) indicate that these organisms produce polycistronic transcripts and that *trans*-splicing is associated with the maturation of their constituent genes [1,3,13,17]. The close linkage of these two schistosome genes and *trans*-splicing of the downstream gene closely resembles the organization of SL2 processed polycistronic transcription units in *C. elegans* [13,17]. Downstream genes in these operons are resolved into individual mRNAs by *trans*-splicing of the specialized SL2 spliced leader at the gene's 5' end and cleavage and polyadenylation at the 3' end of the gene. The first gene in these operons is invariably either not *trans*-spliced or acquires the major SL1 spliced leader. The

average distance between many genes in these operons is ~100 bp, although in several operons it is 300–400 bp [13,17]. Polycistronic transcripts in the kinetoplastida may also have relatively short intergenic regions, e.g. ~100 bp in the tubulin cluster [30]. Based on the precedence in the kinetoplastida and *C. elegans*, the current data are most consistent with the possibility that the detected RNAs represent derivatives of a multicistronic transcription unit. Interestingly, *trans*-splicing of enolase and the close linkage between UbCRBP and enolase are maintained in another trematode, *Fasciola hepatica*, a highly derived digenean taxon. Among trematodes, schistosomes and *Fasciola* are most likely at opposite ends of the digenean radiation suggesting they are highly diverged and that the conservation of gene linkage is significant.

Coupling of polyadenylation and *trans*-splicing has been suggested for processing of polycistronic transcripts in both the kinetoplastida and SL2 genes in *C. elegans* [1,17,30–34]. The exceptionally short 54 nucleotide intergenic region between UbCRBP and enolase suggests that coupling of these two processes might also occur in this locus. Polyadenylation and/or *trans*-splicing within the schistosome UbCRBP/enolase locus may be inefficient or slow compared with *cis*-splicing. This provides one likely explanation for the apparent accumulation of steady state RNAs with the intergenic region present and the upstream UbCRBP and downstream enolase introns removed. Bicistronic cDNAs with exon sequences from two genes connected by an intergenic region representing RNA intermediates for a SL2 processed locus have previously been observed in the *C. elegans* *mai-1/gpd* locus [17].

Using RT-PCR, enolase transcripts which have acquired the SL and retain one or more 5' enolase introns have not been detected. If these steady state RNAs detected represent productive intermediates in mRNA maturation, *cis*-splicing of 5' enolase exons may precede or be required in order for *trans*-splicing to occur. Separate spliceosomes for *cis*- and *trans*-splicing are known to be present in *Ascaris* [35,36], and nearly all *trans*-spliced genes in nematodes and flatworms also contain *cis*-spliced introns. Functional, spatial, or

temporal separation of *cis*- and *trans*-splicing may occur during *trans*-spliced mRNA maturation. This could facilitate and contribute to splice site discrimination in primary transcripts undergoing both *cis*- and *trans*-splicing. cDNAs consistent with *cis*-splicing and polyadenylation preceding *trans*-splicing of *Onchocerca volvulus* actin RNAs have previously been reported [37]. In the schistosome enolase gene, the fidelity of SL addition to the *trans*-splice acceptor site is apparently very high since PCR analyses have not detected SL addition to several internal *cis*-splice acceptor sites. In contrast, similar analyses on several non-*trans*-spliced schistosome genes suggests that internal *cis*-splice acceptor sites in some genes incorrectly acquire the SL and do so at much higher frequencies than in the enolase gene (Davis, unpublished report). Such data might suggest that attributes or processing of *trans*-spliced genes may further contribute to the fidelity of the *trans*-splicing reaction.

Taken together, the following data are consistent with and strongly suggest that UbCRBP and enolase could be expressed as part of a polycistronic transcription unit: (1) the chromosomal organization and unusually short intergenic region between the genes, (2) the structure of the UbCRBP/enolase transcripts, (3) the failure to find transcripts consistent with independent UbCRBP and enolase transcription initiation sites, (4) the conservation of enolase *trans*-splicing and UbCRBP linkage in *Fasciola*, (5) the similarity in organization between UbCRBP/enolase and SL2 *trans*-spliced genes in *C. elegans*, and (6) the compelling data for an association between *trans*-splicing and polycistronic transcripts in the kinetoplastida and SL2 processed genes in *C. elegans*. Enolase *trans*-splicing is conserved in all trematode flatworms examined and in a primitive polyclad flatworm [38,39]. It is tempting to speculate that the conservation of enolase mRNA *trans*-splicing in trematodes and in polyclad turbellarians is associated with maintenance of the organization and expression of the enolase genes within a polycistronic transcription unit.

## Acknowledgements

The authors express deep gratitude to Hardeep Singh for her initial contributions to this work and Tim Nilsen and Chris Tschudi for their comments on the manuscript. The authors also thank Yinghui Cao, Yao Ding, Christian Gorycki, Angelica Parra, Eugenia Ribeiro, and Renee Trudeau for technical assistance. This research was supported in part by a grant from the National Institutes of Health (AI 32709).

## References

- [1] Blumenthal T. *Trans*-splicing and polycistronic transcription in *C. elegans*. Trends Genet 1995;11:132–6.
- [2] Davis RE. Spliced leader RNA *trans*-splicing in metazoa. Parasitol Today 1996;12:33–40.
- [3] Ullu E, Nilsen T. Molecular biology of protozoan and helminth parasites. In: Marr JJ, Muller M, editors. Biochemistry and Molecular Biology of Parasites. New York: Academic Press, 1995:1–17.
- [4] Boothroyd J, Cross G. Transcripts coding for different variant surface glycoproteins in *Trypanosoma brucei* have a short identical exon at their 5' end. Gene 1982;20:279–87.
- [5] Krause M, Hirsh D. A *trans*-spliced leader sequence on actin mRNA in *C. elegans*. Cell 1987;49:753–61.
- [6] Tessier L-H, Keller M, Chan R, Fournier R, Weil JH, Imbault P. Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by *trans*-splicing in *Euglena*. EMBO J 1991;10:2621–5.
- [7] Rajkovic A, Davis RE, Simonsen JN, Rottman FM. A spliced leader is present on a subset of mRNAs from the human parasite *Schistosoma mansoni*. Proc Natl Acad Sci USA 1990;87:8879–83.
- [8] Van der Ploeg L, Liu A, Michels P, De Lange T, Borst P, Majumder H, Weber H, Veeneman G, Van Boom J. RNA splicing is required to make the messenger RNA for a variant surface antigen in trypanosomes. Nucleic Acids Res 1982;10:3591–604.
- [9] Ruppert EE, Barnes RD. Invertebrate Zoology. New York: Saunders, 1994.
- [10] Raff RA, Marshall CR, Turbeville JM. Using DNA sequences to unravel the Cambrian radiation of the animal phyla. Annu Rev Ecol Syst 1994;25:351–75.
- [11] Brusca RC, Brusca GJ. Invertebrates. Sunderland: Sinauer, 1990.
- [12] Nielsen C. Animal Evolution, Interrelationships of the Living Phyla. Oxford: Oxford University Press, 1995.
- [13] Zorio DAR, Cheng NN, Blumenthal T, Spieth J. Operons as a common form of chromosomal organization in *C. elegans*. Nature 1994;372:270–2.

- [14] Maroney PA, Denker JA, Darzynkiewicz E, Laneve R, Nilsen TW. Most mRNAs in the nematode *Ascaris lumbricoides* are *trans*-spliced: a role for spliced leader addition in translational efficiency. *RNA* 1995;1:714–23.
- [15] Nilsen TW. *Trans*-splicing of nematode pre-messenger RNA. *Annu Rev Microbiol* 1993;47:385–411.
- [16] Huang X-Y, Hirsh D. A second *trans*-spliced leader RNA sequence in the nematode *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 1989;86:8640–4.
- [17] Spieth J, Brooke G, Kuersten S, Lea K, Blumenthal T. Operons in *C. elegans*: polycistronic mRNA precursors are processed by *trans*-splicing of SL2 to downstream coding regions. *Cell* 1993;73:521–32.
- [18] Davis RE, Hardwick C, Tavernier P, Hodgson S, Singh H. RNA *trans*-splicing in flatworms: analysis of *trans*-spliced mRNAs and genes in the human parasite *Schistosoma mansoni*. *J Biol Chem* 1995;270:21813–9.
- [19] Davis RE, Davis AH, Carroll SM, Rajkovic A, Rottman FM. Tandemly repeated exons encode 81-base repeats in multiple, developmentally regulated *Schistosoma mansoni* transcripts. *Mol Cell Biol* 1988;8:4745–55.
- [20] Chomczynski P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques* 1993;15:532–7.
- [21] Chomczynski P, Sacchi N. Single-step method of isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;162:156–9.
- [22] Fromont-Racine M, Bertrand E, Pictet R, Grange T. A highly sensitive method for mapping the 5' termini of mRNAs. *Nucleic Acids Res* 1993;21:1683–4.
- [23] Maruyama K, Sugano S. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* 1994;138:171–4.
- [24] Sallie R. Characterization of the extreme 5' ends of RNA molecules by RNA ligation-PCR. *PCR Methods Appl* 1993;3:54–6.
- [25] Volloch V, Schweitzer B, Rits S. Ligation-mediated amplification of RNA from murine erythroid cells reveals a novel class of b globin mRNA with an extended 5'-untranslated region. *Nucleic Acids Res* 1994;22:2507–11.
- [26] Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 1988;85:8998–9002.
- [27] Rother RP. Increasing the specificity of Oligo(dT)-primed PCR. *Biotechniques* 1992;13:524–7.
- [28] Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1989.
- [29] Altschul SF, Boguski MS, Gish W, Wootton JC. Issues in searching molecular sequence databases. *Nat Genet* 1994;6:119–29.
- [30] Matthews KR, Tschudi C, Ullu E. A common pyrimidine-rich motif governs *trans*-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. *Genes Dev* 1994;8:491–501.
- [31] LeBowitz JH, Smith HQ, Rusche L, Beverley S. Coupling of poly(A) site selection and *trans*-splicing in *Leishmania*. *Genes Dev* 1993;7:996–1007.
- [32] Ullu E, Matthews KR, Tschudi C. Temporal order of RNA-processing reactions in trypanosomes: rapid *trans*-splicing precedes polyadenylation of newly synthesized tubulin transcripts. *Mol Cell Biol* 1993;13:720–5.
- [33] Vassella E, Braun R, Roditi I. Control of polyadenylation and alternative splicing of transcripts from adjacent genes in a procyclin expression site: a dual role for polypyrimidine tracts in trypanosomes?. *Nucleic Acids Res* 1994;22:1359–64.
- [34] Kuersten S, Lea K, MacMorris M, Spieth J, Blumenthal T. Relationship between 3' end formation and SL2-specific *trans*-splicing in polycistronic *Caenorhabditis elegans* pre-mRNA processing. *RNA* 1997;3:269–78.
- [35] Maroney PA, YU Y-T, Jankowska M, Nilsen TW. Direct analysis of nematode cis- and *trans*-spliceosomes: A functional role for U5 snRNA in spliced leader addition *trans*-splicing and the identification of novel Sm snRNPs. *RNA* 1996;2:735–45.
- [36] Yu Y-T, Maroney P, Nilsen TW. Functional reconstitution of U6 snRNA in nematode cis- and *trans*-splicing: U6 can serve as both branch acceptor and a 5' exon. *Cell* 1993;75:1049–59.
- [37] Shiwaku K, Donelson JE. *Cis*-splicing and polyadenylation of actin RNA can precede 5' *trans*-splicing in nematodes. *Biochem Biophys Res Commun* 1995;211:49–53.
- [38] Davis RE, Singh H, Botka C, Hardwick C, el Meanawy MA, Villanueva J. RNA *trans*-splicing in *Fasciola hepatica*: Identification of an SL RNA and spliced leader on mRNAs. *J Biol Chem* 1994;269:20026–31.
- [39] Davis RE. Surprising diversity and distribution of SL RNAs in flatworms. *Mol Biochem Parasitol* 1997;87:29–48.

Erratum

Erratum to “Gene linkage and steady state RNAs suggest  
*trans*-splicing may be associated with a polycistronic transcript  
in *Schistosoma mansoni*”  
[Mol. Biochem. Parasitol. 89 (1997) 25–39]<sup>1</sup>

Richard E. Davis \*, Scott Hodgson

*Department of Biological Sciences, Fordham University, Bronx NY, 10458, USA*

---

The publisher regrets that Fig. 1 in the above article was poorly reproduced in the printed issue. The figure is reprinted on the following pages.

---

\* Corresponding author. Tel.: +1 718 8173630; fax: +1 718 8173645; e-mail: rdavis@murray.fordham.edu

<sup>1</sup> PII of original article: S0166-6851(97)00097-2.



**A**U  
b  
c  
r  
b  
p

acgcattaatcagcgacacctcatgattaaagtgtgccaactagtagcaattcagaagtcagaagcgaagagactcgaagatatacttgatagaacatc 100  
gactaatcgagagtgactagctgtttaggaatgtgtggcagcagtgaccactcgttatctgatgagggtgtgtgacagagagccttgggctaagt 200  
attccagtaaaagtaagttaactccaatgtcgaagcattacagaactattggttttgggatttattacggctacagttcaaaagtcacctatt 300  
gcctacactaacctattgttatccagttataaaacggtccaaacgaagtcgaatccgaccacactgccattatttcaagaaccacatcacaaaagt 400  
atgattccactcatattataatattcactaggtcggtggcactaacttgaacagagggatttcatgtatttagaatacgaaggattcttctcgaagg 500

agtttctaggtcgctgctctattttcggttttgaatatagattgaaccataattgttttACTGTATAAATGATGACCATGTAGATCGATTGGTAAA 600

**UbCRBP Exon 1 Alternative Splice Donor Sites**

A B C D  
G C A G T A A G T C A C G T A G G T T A T T C G G G A C A T G G G T C G G T T G G G A G T T T G g t a t g c t c a a t c a t t t a t t t t c a g t t a c c a c c a t t a t t a c a a a t c t a g t t c g 700

gcagctctggactgtagggtattcattattctagtttaattgtattccagctctaaattctatacaaccacagaacccaaactgcggaataaaagaagcg 800  
acaaagtaattaaaaaggaattgtcgatttaccctcagatgttagttcttcaagaaaaatggtaattctcccccttcgactatttaaggaagattgata 900  
gatataatttctgtgaagtaccagctcgaatgtgtgagatgttaacatacactcgcaaaaatctaccatacacacccaagggtttatggcccagaaactact 1000  
aaaatagtttgaagcggtaattggacaacaggtgcttccacttaacgcgacttatctgtcaccatgctcccaattagggtcttccataatttagattt 1100  
aatgtacccgttttacttaattcggatttaagcagtttaatttgaacggatataagatgtactatgcatgttgctcggttgatataaagtagttaccag 1200  
atagtcaggcgcttattcctcttataagtagcaggttaagtagtattgttagaaaaaccgttttcatgaacgctattatgtacatttttaggtctagagtgat 1300  
ataacgatctccacaagttataaaagcactgtaacacatgttagtcaatattgtcgaagtagtgcacgcaatataatagatatcactcggttaacttagtc 1400

tttctgtgtaaaaaatccgttttcatataagttcaaaaatatttcatataagatcctattcattaccttacagGCTAGCGAAATATATTCGGGATGGCAT 1500  
MetAlaT

ACCTCCGAAGTGGTGGGAATTTATAGCATCGTATCGTTCAGTCTTAATCGAATGAGTGATCTTCAACAAAAGATAAAGATGTTCCGATTCAACAATGC 1600  
yrLeuArgThrGlyAlaGluPheIleAlaSerTyrArgSerSerLeuAsnArgMetSerAspLeuGlnGlnLysIleLysMetPheGlyPheAsnAl

TTACTATAATCAGGTCGGACTTTTGAACATGATTGGCTGCCTCATAGTCTGTTGGCTCGAAGCTCTACGTCGACTTCCAGTGAACACAGGAAGCT 1700  
aTyrTyrAsnGlnValGlyLeuLeuLysHisAspTrpLeuProHisSerProValTrpLeuGluAlaLeuArgArgLeuProArgGluLeuGlnGluAla

AGGGACTTTCGAATCGCAGCAGCGCTCGCTTCTCTACGCTTCGAACATGTTTGGCCGAAGAACAGTGGACTACTATTGAAGATGATATCCCACTACTGG 1800  
ArgAspPheArgIleAlaArgAlaSerLeuLeuTyrAlaSerLysHisValLeuProLysGluGlnTrpThrThrIleGluAspAspIleProTyrLeuG

AGCCGTACGTTAACGTTGTTGTGAAGAAATGAGTGACCGGTCAAATTGGGATAATTTTGTCAATCCTGAGATATATAGTGAATAAATGTAATGTTTAC 1900  
luProTyrValAsnValValLysGluMetSerAspArgSerAsnTrpAspAsnPheValAsnProGluIleTyrSerGlu

**UbCRBP Polyadenylation Signal** **Enolase Trans-splice Acceptor**

ATTGTCTGTTAGTGTCAAATTAGTCTAATAAAAGTTGTTGAAAAGCAttgttcttactgtcctatgttctaacctagctaaaaacacctttttcttcgcag 2000

**Intergenic Region**

AAAAATGTCATTTTAAACGATCCACGCTCGTCAAATTTTCGATAGTCGGGGGAATCCACAGTTGAGGTGGATTGAAAAACATCTAAAGgtatgttggtag 2100  
MetSerIleLeuThrIleHisAlaArgGlnIlePheAspSerArgGlyAsnProThrValGluValAspLeuLysThrSerLysG

acaataagtttatgaacctttgtcttcttatttccacttggtagacacaaattttcacatgtactaatgtgatctcaccactgtcgttgtagactagtc 2200  
gatagtagcttgggtttttagcttaccacaaatttgcctgttagtgggtttgtcaagggttaataactcagtgagatagatatcatcctaatac 2300  
ttattgttaatttgcaggtgttgcgttagcggattttaaagttccagtttaagtttaactttgataataatcttatgtatagattcaaatgtaaaagtc 2400

ttttttgtccagGTTTGTTCGCGCTGCCGTTCCAAGTGGAGCTTCGACAGGTGTCATGAGGCTCTTGAGTTACGTGATACAAATCAAAGGCTTATA 2500  
lyLeuPheArgAlaAlaValProSerGlyAlaSerThrGlyValHisGluAlaLeuGluLeuArgAspThrAsnSerLysAlaTyrM

TGAAGAAAGGGTTTTAAACAGCAGTCAGTAATGTGAATAAAATTATTGCTCCAGCTCTAATAACAAAAATATCCCGTAACAAATCAGGCAGCTATTGA 2600  
etLysLysGlyValLeuThrAlaValSerAsnValAsnLysIleIleAlaProAlaLeuIleAsnLysAsnIleProValThrAsnGlnAlaAlaIleAs

CAATATATGATTGATCTTGATGGTACGGAAAAACAAGGtgagtttaattgatgacttatatcctcgctacataagttgtggataaagttattgttagat 2700  
pLysTyrMetIleAspLeuAspGlyThrGluAsnLys

**B**U  
b  
c  
r  
b  
p

...TCTGCGAGCAAAACAGTGCTTCCAAAAGAGCAGTGGACTGCCATCGAAGACGATATTCCCTATTTGGATCCATACATCGAAGTGGCAAAAAGGAAT 100  
...SerAlaSerLysThrValLeuProLysGluGlnTrpThrAlaIleGluAspAspIleProTyrLeuAspProTyrIleGluValAlaLysLysGluT

GGAAGAGAAAGCGGAGTGGGATCATTTTGTGAACCCGGAACATATCCGtgataattgcttgttaatactgtagaatttaatacatagtgccctaaag 200  
rplysGluLysAlaGluTrpAspHisPheValAsnProGluThrTyrPro

**Fasciola Enolase Trans-splice Acceptor**

agtttgcctttaaagtgaaatgtgttcgctgttctaacacgtctcctctggtttccagAACATGGCAATAAAAGCGATCCACGCCCCGACAAATAT... 300  
MetAlaIleLysAlaIleHisAlaArgGlnIleP...

E  
n  
o  
l  
a  
s  
e

Fig. 1.



Fig. 1. Closely linked gene upstream of enolase encodes UbCRBP. (A) *Schistosoma mansoni* UbCRBP gene sequence extending into exon 3 of the downstream enolase gene (exon 1 of the enolase gene is the SL). Lower case letters represent upstream regions, introns, or the intergenic region between UbCRBP and enolase. Uppercase letters represent exon sequences (light shading, UbCRBP; darker shading, enolase) and the predicted translation products of the encoded mRNAs are shown below the exon sequence. The putative polyadenylation signal, intergenic region, and *trans*-splice acceptor site for enolase are marked. The alternate splice donor sites contributing to the definition of UbCRBP exon 1 are shown with triangles with the GT dinucleotide underlined. Polyadenylation sites determined by 3' RACE are marked with diamonds and the number of clones identified corresponding to each site are provided in parentheses. (B) Linkage between UbCRBP and the *trans*-spliced enolase gene in *Fasciola hepatica*. Only the 3' end of *Fasciola* UbCRBP (accession No. U91748) and a portion of enolase exon 2 are shown (exon 1 is the *Fasciola* spliced leader [37]). Letters and shading are the same as in (A). (C) Multiple alignment of the upstream schistosome and *Fasciola* ORFs with phylogenetically diverse UbCRBP proteins. Dashes indicate gaps used to generate the optimal alignment, dark shading indicates identical residues, and light shading indicates conservative amino acid substitutions.