The EMBO Journal vol.7 no.5 pp.1431 – 1437, 1988

Trypanosome *trans*-splicing utilizes 2′ – 5′ branches and a corresponding debranching activity

Richard E.Sutton and John C.Boothroyd

Department of Medical Microbiology, Sherman Fairchild Building, Standford University School of Medicine, Stanford, CA 94305, USA

Communicated by B.Stocker

The 5′ ends of trypanosome mRNAs consist of an identical sequence of 39 nucleotides which is derived from a discrete transcript of ~140 nucleotides (medRNA). It has been proposed that generation of chimeric mRNAs in trypanosomes occurs by the process of *trans*-splicing involving medRNA and an acceptor RNA. Part of the basis for this suggestion comes from the ability of HeLa cell extracts (known to contain debranching activity) to catalyze the release of the intron portion of medRNA (minRNA) implying a Y-branched intermediate in the splicing process. Here we provide direct chemical analysis that miniRNA is attached to higher mol. wt RNA molecules by a 2′–5′ phosphodiester bond (i.e. as a branched structure). We also demonstrate that trypanosomes have substantial amounts of debranching activity which is similar in nature to that of HeLa cells. These results provide further evidence for *trans*-splicing in trypanosomes and highlights its similarity to *cis*-splicing in other eukaryotes.

Key words: trypanosomes/splicing/branch/debranching

Introduction

All examined mRNAs of *Trypanosoma brucei*, including those of the variant surface glycoproteins, have precisely the same 39 nucleotides at their 5′ ends (Boothroyd and Cross, 1982; Van der Ploeg *et al.*, 1982; De Lange *et al.*, 1984; Dorfman and Donelson, 1984; Parsons *et al.*, 1984). The mini-exon encoding this untranslated leader is part of a 1.35-kb repeat which is tamdemly repeated ~200 times (in one or more clusters) in the trypanosome genome (De Lange *et al.*, 1983; Michiels *et al.*, 1983; Nelson *et al.*, 1983). Homologous sequences have been found on the mRNAs and in the genomes of related kinetoplastids (Nelson *et al.*, 1984; De Lange *et al.*, 1984; Muhich *et al.*, 1987), and analogous (but not homologous) sequences have been recently detected in the metazoan *Caenorhabditis elegans* (Krause and Hirsch, 1987).

The 1.35-kb mini-exons repeat in *T.brucei* directs the synthesis of a short, capped transcript (~140 nt; Freistadt *et al.*, 1987; Perry *et al.*, 1987; Sutton and Boothroyd, 1988) termed the mini-exon-dervied RNA (medRNA) or spliced-leader RNA (Campbell *et al.*, 1984; Kooter *et al.*, 1984; Milhausen *et al.*, 1984). The 5′-most 39 nucleotides of medRNA correspond to the mini-exon itself, and this portion is immediately followed by a consensus 5′ splice-site signal (GUAUGA). Immediately upstream of all protein-coding exons thus far characterized is the consensus 3′ splice-site signal $[C/U]_nNNAG$ (Borst, 1986). These observations, together with the fact that the mini-exon repeats and some structural genes are on different chromosomes (Van der

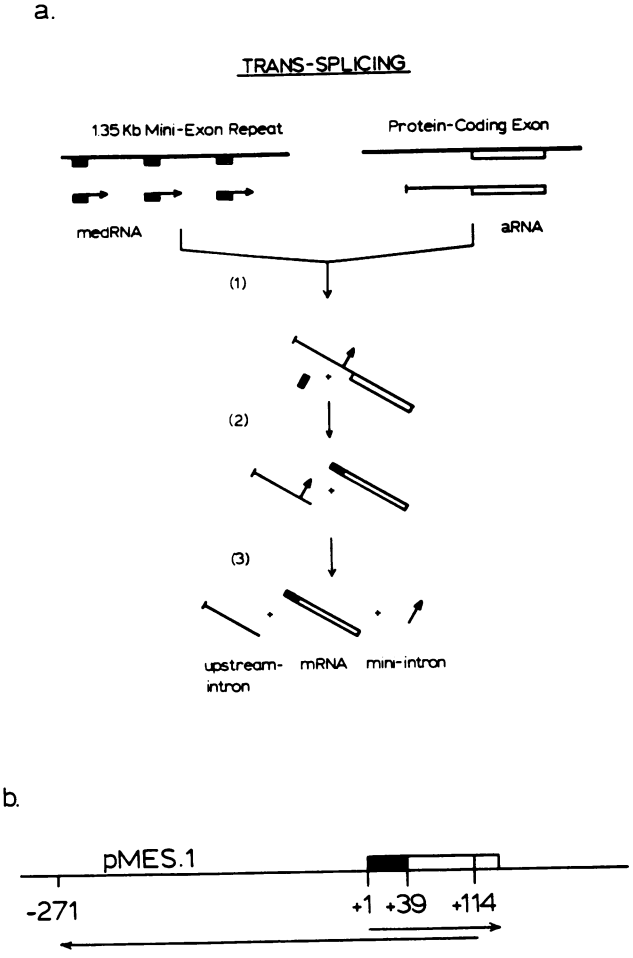

Fig. 1. Proposed model of *trans*-splicing in trypanosomes. The 1.35-kb mini-exon repeats direct the synthesis of medRNA; the protein-coding exon directs the transcription of the acceptor RNA (aRNA). The first step in *trans*-splicing involves formation of a branched species [comprising aRNA and the intron portion of medRNA (minRNA)] and free mini-exon. In the second step, the two exons are joined so that the two introns are released as a discrete, branched species. The third step comprises the debranching of the branched species into the two component introns. In this reaction, there are three final products: the upstream intron of aRNA, minRNA and the mature mRNA (reprinted, with permission, from Sutton and Boothroyd, 1986). Below the model is a schematic map of a portion of a cloned mini-exon repeat (pMES.1, see Campbell *et al.*, 1984) showing the mini-exon (solid box), mini-intron (open box), medRNA transcript (rightward arrow) and the protecting anti-sense RNA probe (leftward arrow). Nucleotides are numbered from the 5′ end of medRNA. Note the revised numbering based on recent results indicating the mini-exon is 39 nucleotides long rather than 35 as originally reported (Freistadt *et al.*, 1987; Perry *et al.*, 1987).

Ploeg *et al.*, 1984; Guyaux *et al.*, 1985), led to the hypothesis that mRNAs are generated by the process of 'discontinuous transcription'. In other words, a chimeric mRNA is produced by transcription of two unlinked loci.

Recent evidence suggests that the trypanosome generates these chimeric mRNAs by the process of *trans*-splicing (Murphy *et al.*, 1986; Sutton and Boothroyd, 1986; Laird *et al.*, 1987). The downstream portion of medRNA (corresponding to nucleotides +40 to +140, and termed mini-intron or minRNA) exists at significant levels in trypanosome RNA (Sutton and Boothroyd, 1986) and can be released as a discrete species from polyadenylated RNA by HeLa cell extracts containing debranching activity. Evidence so far published suggests that this *trans*-splicing reaction is directly analogous to *cis*-splicing of mRNAs in higher eukaryotes (see Figure 1a), except that instead of a lariat, the trypanosome intermediate is probably a Y-structure.

Here we present direct chemical evidence that minRNA can exist as part of a branched series through a 2′−5′ phosphodiester bond, the central nucleotide of the branch being adenosine. We also demonstrate that trypanosomes have substantial amounts of debranching enzyme activity, and that this activity has similar properties to that of HeLa cells. These results provide additional evidence in support of *trans*-splicing of mRNAs in trypanosomes and further illustrate the similarity to *cis*-splicing of higher organisms. The evolutionary implications of this are discussed.

Results

Incorporation of radioactive phosphate into trypanosome RNA

After a 4-h labeling of $\sim 10^9$ trypanosomes in 25 ml of low phosphate media, ~15% of the total input $[^{32}P]H_3PO_4$ was incorporated into RNase-sensitive, TCA-precipitable nucleic acid. Of this, after two rounds of oligo(dT)−cellulose selection, 98−99% remained as poly(A)$^-$ RNA (4−5 mCi). During these manipulations, it was important to minimize the time that the RNA was precipitated in pellet form to prevent autolysis due to the very large amounts of ^{32}P involved.

Isolation of branched RNA molecules

The uniformly labeled poly(A)$^-$ RNA described above was first RNase-protected with non-radioactive, anti-sense RNA probe extending from −271 to +114 of the mini-exon repeat (see Figure 1B). As shown in Figure 2a, two protected species are observed. The one at ~114 nucleotides represents the fragment of medRNA extending from +1 to +114, and the one at ~75 nucleotides represents minRNA from +40 to +114, probably in both branched and unbranched configurations. The diffuse but strong signal just below this is of unknown origin but seems likely to be an artefact of the gel electrophoresis and large amounts of material being loaded as most of this label was readily dialyzed out. The similar levels of the 114-mer and 75-mer are surprising given previous determinations by Northern blot analysis of the relative amounts of medRNA and minRNA (free and in branches). This discrepancy is reproducible and presumably reflects limitations and differences between detection methods and/or possible differences in the higher order structure of the two molecules.

After electroelution and dialysis of the RNA species of ~75 nucleotides, it was separately treated with either nuclease P1 or the combination of RNase A, T1 and T2 and chromatographed on two-dimensional cellulose F TLC plates (Figure 2b and c respectively). It is clear that both treatments result in multiple radioactive spots. Four are easily identifiable (based on amounts, known migrations and co-chromatographed standards) as the four nucleoside mono-

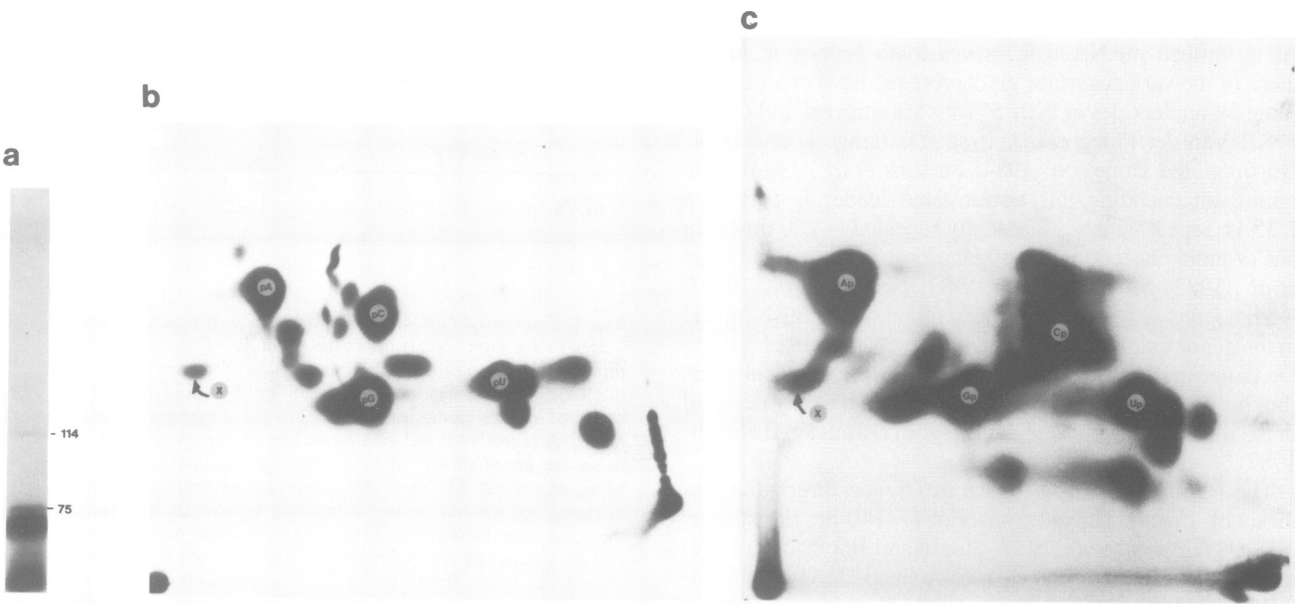

Fig. 2. Preparation of core branch structure. (**a**) RNase protection of minRNA. Using the non-radioactive, anti-sense RNA probe of Figure 1, uniformly labeled poly(A)$^-$ RNA was RNAse-protected (Myers *et al.*, 1985) and the product electrophoresed on a 7 M urea−5% polyacrylamide gel. Markers were 3′-end labeled *Msp*I fragments of pAT153 (not shown). (**b**) Nuclease P1 digest of minRNA. The band of ~75 nucleotides of Figure 2a was electroeluted, dialyzed against a large volume of water and dried under vacuum. A portion was then treated with nuclease P1 and chromatographed on two-dimensional cellulose F TLC plates (see Materials and methods). The first dimension is vertical, the second dimension is horizontal. Migration of the 5′ NMP markers is shown. The species marked with an 'X' was that ultimately found to be the core branch. (**c**) RNase A, T1, T2 digest of minRNA. As for Figure 2b except treatment was with these three RNases, and markers are the 3′ NMPs.

phosphates (NMPs). The minor spots of Figure 2c were scraped and eluted, and multiple enzymatic digests were performed. The specificities of the various treatments on a core branch are shown beside Figure 3b. Of ten spots analyzed, only one spot (marked 'X' in Figure 2c) had the expected characteristic of a core branch, i.e. of having the same profile after both alkaline phosphatase and nuclease P1 treatment, with conversion of a significant amount of ^{32}P to orthophosphate (Figure 3a–c). Furthermore, the original mobility of this spot is approximately as previously reported for the core branch trinucleotide in this chromatographic system (Konarska *et al.*, 1985). It should be noted that this spot was not generated when the radioactive smear just below the 75 nucleotide species of Figure 2a was treated with the three RNases, indicating that it is not a contaminant from this uncharacterized material (data not shown).

To confirm further that this spot had the structure of a trinucleotide core branch (two 3′ nucleoside monophosphates connected via 5′ phosphodiester bonds to the 2′ and 3′ positions respectively of a third nucleotide), the sample was treated with nuclease P1 to remove the 3′ phosphates, phenol/chloroform extracted, and then oxidized with periodate followed by β-elimination. The resulting nucleotide was desalted and chromatographed on two-dimensional cellulose F TLC plates. The resulting autoradiogram (Figure 3d) gave rise to a single major spot which migrated adjacent to the 3′,5′ ADP standard which was co-chromatographed, consistent with it being the isomeric 2′,3′ ADP nucleotide. In other experiments we have determined that this is indeed the mobility of 2′,3′ ADP relative to 3′,5′ ADP (data not shown). These results indicate that the core branch being analyzed has an adenosine residue at the branch point.

The putative branched nucleotide which resulted from the nuclease P1 digest (marked with an 'X' in Figure 2b) was also subjected to periodate oxidation and β-elimination. After desalting, the sample was chromatographed on cellulose F TLC plates, and the result was identical to that of Figure 3d (data not shown).

Characterization of the nucleotide at the 2′ position of the branch

In order to determine the first nucleotide of minRNA which is covalently linked to the branched structure, total trypanosome RNA was either treated or not with HeLa debranching enzyme. The resultant RNA was treated with alkaline phosphatase followed by polynucleotide kinase using [γ-^{32}P]-ATP, and the RNA was hybridized with non-radioactive, anti-sense RNA probe-spanning nucleotides −271 to +114 of the mini-exon repeat (Figure 1b). Following RNase digestion, protected RNA was extracted and electrophoresed on a denaturing polyacrylamide gel. The only observable increase in amount of radioactive RNA upon debranching was at ~75 nucleotides; control experiments lacking phosphatase treatment resulted in virtually no signal (data not shown). The RNA species of 75 nucleotides with and without debranching treatment was electroeluted, dialyzed and treated with nuclease P1. The resultant two-dimensional cellulose F TLC plates (Figure 4a and b) indicate that debranching leads to a specific increase in the amount of 5′ GMP. This result is consistent with branched minRNA having a guanosine at the 2′ position of the branch. The background levels of 5′ GMP and 5′ AMP in the non-debranched material (Figure

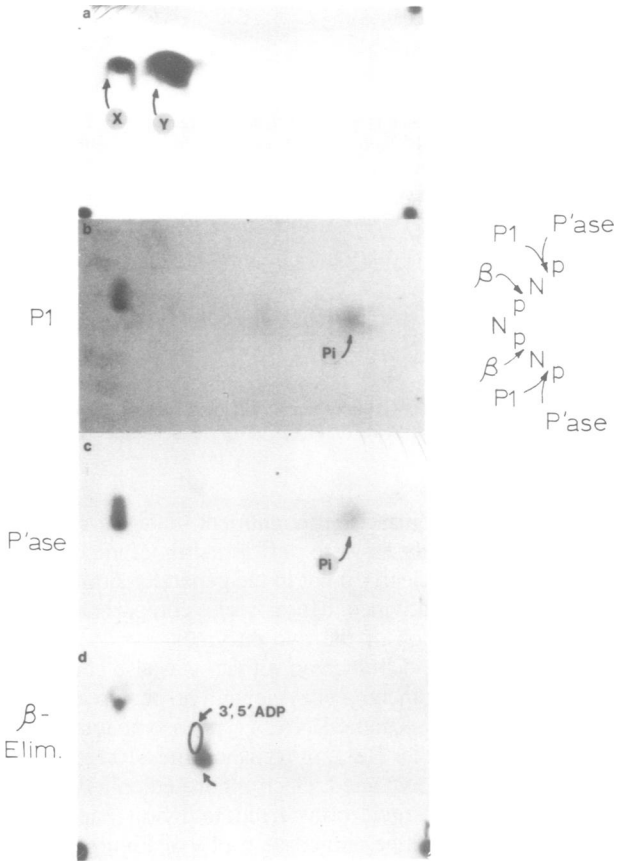

Fig. 3. Identification of the branched oligonucleotide. The 'X' species of Figure 2c was scraped, eluted with water, subjected to enzymatic and chemical treatments and chromatographed on cellulose TLC plates as in Figure 2. (**a**) Untreated. The 'X' species of Figure 2c is indicated. The spot marked by 'Y' is a second species, separately eluted from the plate shown in Figure 2c. It was analyzed on this control plate only and is irrelevant to the study being reported here. (**d**) Treatment with nuclease P1. Shown to the right are the expected cleavage sites of the various treatments on the core trinucleotide branch. (**c**) Treatment with alkaline phosphatase. (**d**) First treated with nuclease P1 (to remove the 3′ phosphates), and then oxidized and β-eliminated (to remove the 2′ and 3′ bases). The circle indicates the mobility of co-chromatographed 3′,5′ ADP. Some of the original, untreated species is visible, indicating incomplete digestion by nuclease P1. The unlabeled arrow indicates the putative 2′,3′ ADP (see text).

Fig. 4. Identification of the first nucleotide of minRNA in branched configuration. Total trypanosome RNA was either treated or not with HeLa debranching enzyme followed by alkaline phosphatase and subsequent 5′ end-labeling with polynucleotide kinase. The RNA samples were then hybridized with non-radioactive, anti-sense RNA probe of Figure 1 followed by treatment with RNase. Protected speci were electrophoresed on a 7 M urea−5% polyacrylamide gel, and in each case, the species of ~75 nt was electroeluted, dialyzed against water, dried, treated with nuclease P1 and chromatographed on two-dimensional TLC plates on Figure 2, along with the four cold 5′ NMPs. Mobility of the relevant co-chromatographed markers is shown. (**a**) No pretreatment. (**b**) Pretreatment with HeLa cell S100 extract containing the debranching activity.

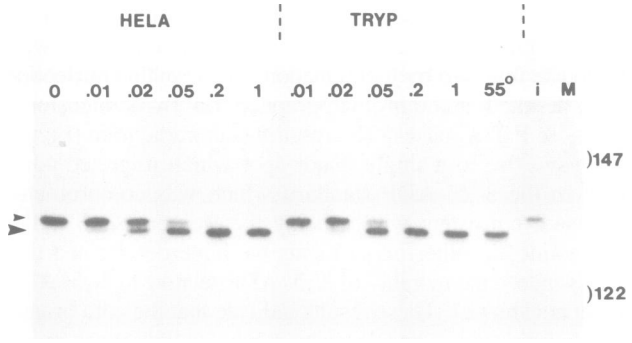

Fig. 5. Titration of trypanosome and HeLa S100 extracts containing debranching enzyme. Trypanosome S100 extract was prepared as described in Materials and methods and HeLa S100 extract was a gift of the laboratory of Dr M.R.Green. The substrate used here is lariat intron 1 of the human β-globin gene, isolated from a two-dimensional denaturing polyacrylamide gel. Lariat RNA was debranched as previously described (Ruskin and Green, 1985), extracted (Sutton and Boothroyd, 1986), and separated on a 7 M urea−5% polyacrylamide gel. The dilutions of both extracts are indicated at the top of each lane (i.e. 0.05 indicates 0.125 μl relative to the strandard 2.5 μl was used) and the lariat and linear forms are indicated at the left of the autoradiograph with small and large arrowheads respectively. Input (untreated) lariant is in **lane i** and markers are in **lane M**. The lane marked 55° was a preliminary assessment of the heat stability of the trypanosome extract under standard reaction conditions.

4a) are of unknown origin although the 5′ GMP may be derived from minRNA free within the cell (Sutton and Boothroyd, 1986).

Detection of debranching activity in trypanosome extracts

Trypanosome S100 extract was tested for its ability to debranch lariat RNA. The substrate used was the lariat intron from the first intron of the human β-globin gene, kindly provided by Phillip Zamore of Dr Michael Green's laboratory (Harvard University). The trypanosome extract was found to contain significant amounts of debranching activity. The extract was titrated and compared with the amount of debranching activity from similarly prepared HeLa cell S100 extracts (Figure 5). Given the fact that the protein concentration of the HeLa extract is approximately the same as that of the trypanosome extract (10 mg/ml), both S100s have a similar specific activity. As well, note that the trypanosome extract is virtually devoid of other RNase activity.

Properties of the trypanosome debranching enzyme

The optimum pH of the trypanosome debranching activity was compared with that of the HeLa enzyme (Figure 6a). The HeLa enzyme has a broad pH optimum centered around pH 8.0, whereas the trypanosome enzyme has a more pronounced dependence upon pH, its optimum being near pH 7.6. Both enzymes are inhibited by high concentrations of KCl (Figure 6b): the trypanosome enzyme appears to be completely inhibited at 100 mM KCl; the mammalian enzyme is fully inhibited at 200 mM KCl. Ruskin and Green (1985) have reported that the HeLa enzyme loses activity after a 15-min heat pretreatment at 52°C. We directly compared both enzymes by pretreating with heat (Figure 6c). The trypanosome enzyme is only slightly more heat stable than the HeLa enzyme. As expected, both enzymes were

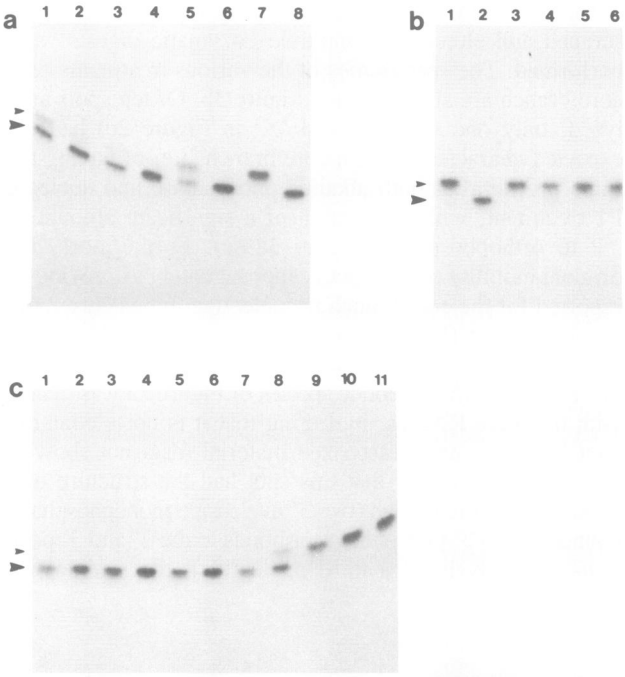

Fig. 6. Characteristics of trypanosome debranching enzyme. Lariat RNA was the same as in Figure 5 and, except as noted below, standard debranching reactions were performed (Ruskin and Green, 1985). Extraction and electrophoretic analysis of the reactions is as in Figure 5. Odd-numbered lanes are treatment with trypanosome and even-numbered lanes are treatment with HeLa S100 extract. Other symbols as in Figure 5. **(a)** pH dependence. **Lanes 1** and **2**, 20 mM Hepes, pH 7.0; **lanes 3** and **4**, 20 mM Hepes, pH 7.6; **lanes 5** and **6**, 20 mM Hepes, pH 7.9; and **lanes 7** and **8**, 20 mM EPPS, pH 8.5. **(b)** KCl dependence. **Lanes 1** and **2,** 100 mM KCl; **lanes 3** and **4, 200 mM KCl; lanes 5** and **6,** 300 mM KCl. **(c)** Heat pretreatment for 15 min. **Lanes 1** and **2,** 45°C; **lanes 3** and **4,** 50°C; **lanes 5** and **6,** 55°C; **lanes 7** and **8,** 60°C; **lanes 9** and **10,** 65°C. **Lane 11** is input lariat RNA. **Lanes 3, 4, 7, 8:** only 0.25 μl of extract; **lanes 1, 2, 5, 6:** 2.5 μl of extract.

susceptible to proteinase K pretreatment (data not shown). Although these analyses were performed on crude enzyme preparations, they clearly point to the generally similar properties of the two activities. More precise comparisons must await the purification of the two enzymes.

The trypanosome debranching activity was also compared with the HeLa debranching enzyme with respect to its ability to debranch trypanosome RNA. Trypanosome total RNA was treated with either HeLa or trypanosome S100 extracts, phenol/chloroform extracted, electrophoresed on a denaturing polyacrylamide gel, transferred to nylon paper and probed with radioactive anti-sense probe of Figure 1. The resulting Northern blot shows that, as expected, the trypanosome extract is also capable of releasing minRNA from branched molecules (Figure 7).

Discussion

We demonstrate here by chemical and enzymatic analyses independent of debranching enzyme that minRNA of trypanosomes is covalently attached to other RNA species by a 2′−5′ phosphodiester bond. The nucleotide at which this branching occurs appears to be adenosine. The nucleotide in 3′−5′ linkage to the adenosine was not identified because only a limited amount of material was available.

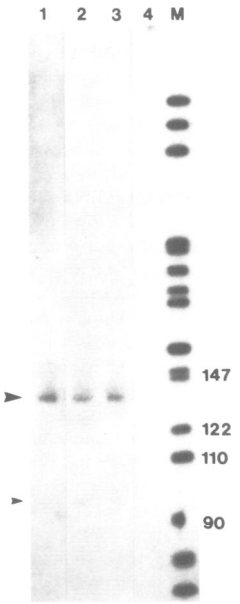

Fig. 7. Treatment of trypanosome RNA with trypanosome and HeLa S100 extract. Total trypanosome RNA was prepared from DEAE-Sephacel purified trypanosomes by the hot phenol method (Boothroyd and Cross, 1982), treated with either HeLa or trypanosome S100, extracted, electrophoresed on a 7 M urea−5% polyacrylamide gel, transferred to genetrans paper (Plasco, Woburn, MA) and hybridized with uniformly radioactive probe of Figure 1. Hybridization and washing conditions were as previously described (Sutton and Boothroyd, 1986). **Lane 1**, mock-treated RNA; **lane 2**, treated with HeLa S100; **lane 3**, treated with trypanosome S100; **lane 4**, yeast transfer RNA (Sigma Chemicals) treated with trypanosome S100. Markers (M) are 3′ end-labeled *Msp*I fragments of pAT153.

These results confirm and extend the work of previous studies in which debranching enzyme of HeLa cells was used to demonstrate that minRNA is attached to both high and low mol. wt trypanosome RNA by a covalent bond (Murphy *et al.*, 1986; Sutton and Boothroyd, 1986; Laird *et al.*, 1987). The fact that branched RNA species occur in poly(A)$^-$ RNA is not surprising since branched intron molecules which have been released from the protein-coding exons should not be polyadenylated (Figure 1).

We have shown that one of, if not the, major core branch in trypanosome *trans*-splicing has a central adenosine. Unfortunately, we cannot conclude that this is the only nucleotide to serve this function; although unlikely, others may have co-migrated with the four NMPs (which were not analyzed further), and there were several, very minor spots whose small amounts precluded analysis. The use of adenosine in branch formation is consistent with the results obtained *in vivo* and *in vitro* in HeLa cells and in yeast (for review, see Green, 1986; Padgett *et al.*, 1986). In yeast, branching probably occurs exclusively at adenosine residues (UACUAAC; A being the branch point), whereas in HeLa cells branching may occur to a minor extent at cytidine residues as well (Wallace and Edmonds, 1983). Experiments in yeast *in vitro* and *in vivo* (Langford *et al.*, 1984; Newman *et al.*, 1985; Jacquier and Rosbash, 1986) and in HeLa cells *in vitro* (Hornig *et al.*, 1986) using mutated templates suggest that branch formation occurs at adenosine much more readily than cytidine residues. The molecular basis for this preference is unknown, but it was clearly established early in the evolution of nuclear pre-mRNA splicing. Interestingly, adenosine is also used in the self-splicing of class II mitochondrial introns (Peebles *et al.*, 1986; van der Veen *et al.*, 1986).

Using HeLa debranching enzyme, we also show that branches which contain minRNA have a guanosine at the 2′ position. Preliminary experiments using other enzymatic digests indicated that the sequence of nucleotides at the 2′ position of the branch is 5′-GpU (data not shown), indicating that, as expected, no extra nucleotides are added to minRNA during the process of *trans*-splicing.

The precise locations of branch sites within the introns of trypanosome acceptor RNAs have yet to be determined. Such determinations should provide confirmatory data for the general use of adenosine as the branched nucleotide. Examination of the sequences of putative acceptor RNAs indicates that in all cases except HSP 70 (Glass *et al.*, 1986) there is a consensus branch site sequence (YYRAY, where Y is pyrimidine and R is purine) within 50 nucleotides of the 3′ splice site (Borst, 1986). Heat shock disrupts mRNA splicing in both higher eukaryotes (Yost and Lindquist, 1986) and trypanosomes (M.L.Muhich and J.C.Boothroyd, submitted) yet the trypanosome HSP 70 mRNA is synthesized at increased levels and accurately *trans*-spliced (M.L.Muhich, M.Hsu and J.C.Boothroyd, unpublished results). Whether this differential sensitivity to heat shock is partially due to a difference in the branch point remains to be determined.

We also demonstrate here that trypanosomes have substantial amounts of debranching activity with similar properties to that already characterized from mammalian cells (Ruskin and Green, 1985; Arenas and Hurwitz, 1987). In other experiments, we have been able to show that, like the HeLa enzyme, the trypanosome activity also leaves a 5′ monophosphate after debranching of the 2′−5′ phosphosdiester bond (data not shown). It is interesting to note that the specific activities (based on per mg protein) of the two enzymes in crude preparations are roughly similar, and that each enzyme is active against the heterologous branch.

The precise substrate or sequence requirements of each enzyme have not been fully addressed. Although no trypanosome *cis*-introns have been described, if they do exist, the activity described here would appear capable of debranching them. Alternatively, as neither the trypanosome nor HeLa activity has been purified to homogeneity, we cannot exclude the possibility that these extracts contain two debranching enzymes, one for *cis*- and one for *trans*-spliced products (lariats and Y-branches respectively). This important point needs resolution.

The physiological role(s) of the debranching enzyme is not known but seems most likely to enable the complete degradation of lariat or branched molecules (no other non-vesicular intracellular nuclease is known to attack 2′−5′ bonds, whether in branched linkage or not). Debranching probably occurs before the RNA species is degraded down to the core (trinucleotide) branch since this is a poor substrate for the debranching activity (Ruskin and Green, 1985; R.E.Sutton, unpublished observations). It is most likely that the free minRNA previously observed in trypanosome RNA preparations is due to the action of debranching enzyme upon branched molecules (Sutton and Boothroyd, 1986), although it has also been suggested to be due to conventional endoribonuclease cleavage of a hyper-sensitive site in medRNA (Murphy *et al.*, 1986; Laird *et al.*, 1987).

Given the fact that the HeLa and trypanosome enzyme

have similar activities and other properties, it seems likely that the primary sequence if not the secondary structure of the enzyme has been conserved throughout evolution of the eukaryotes. Once either gene is cloned, this can be readily tested.

Previous work has already made clear the great similarity of trypanosome *trans*-splicing and conventional *cis*-splicing (e.g. the conservation of splice-site sequences). The fact that trypanosomes utilize 'conventionally' branched RNA (albeit in an open, rather than closed, lariat configuration) and a corresponding debranching enzyme extends this striking similarity. As trypanosomes apparently diverged very early from the main eukaryotic lineage (Sogin *et al.*, 1986) it is clear that many of the details of nuclear pre-mRNA splicing are very ancient. The exact evolutionary relationship between these two types of splicing must now await further work on their mechanistic details. Of particular interest will be the analyses of *Caenorhabditis elegans* in which both types of splicing may occur (Krause and Hirsch, 1897). This may further illuminate the role and relationship of these two processes.

Materials and methods

Reagents

RNase A and T2, sodium periodate, biological buffer EPPS (*N*-[2-hydroxyethyl]-piperazine-*N'*-3-propane-sulfonic acid) and nuclease P1 were from Sigma Chemicals; dimethylsulfide (DMS) and isobutyric acid were from Aldrich Chemicals. RNase T1 was from Bethesda Research Laboratories, and 3,5′ ADP was from Pharmacia Fine Chemicals. The standard 2′,3′ ADP was produced by treating the tetraribonucleotide branch pGA^G_C (i.e. adenosine branched 2′,5′ to G and 3′,5′ to C) (Kierzek *et al.*, 1986; a gift of Dr Mary Edmonds, Unviersity of Pittsburgh) with RNase T1, then performing a $NaIO_4$ oxidation and β-elimination (see below) to leave 2′,3′ ADP. Triethylamine-bicarbonate (TEAB) was made by bubbling CO_2 (from dry ice) through a 1 M solution of redistilled triethylamine until the pH reached ~7.0. This solution was stored at +4°C. RNase-free calf intestinal alkaline phosphatase was purchased from Boehringer-Mannheim. Lariat RNA from human β-globin intron 1 (the 143-nucleotide RNA species of Ruskin and Green, 1985) and HeLa cell S100 extract were gifts of the laboratory of Dr M.R.Green (Department of Biochemistry and Molecular Biology, Harvard University). TLC plates (Cellulose F, 0.1 mm glass-backed with fluorescent indicator) were from Merck and Co., Inc. Saturated $(NH_4)_2SO_4$ was prepared by dissolving 900 g in 1 l water by heat, filtering using a Buchner funnel, cooling to room temperature and adjusting the pH to 7.8 with NH_4OH. Trypanosome S100 extract was prepared by first purifying bloodstream form trypanosomes using a DEAE-Sephacel column as described by Lanham and Godfrey (1970). Cells were washed in Separation buffer (44 mM NaCl, 57 mM Na_2HPO_4, 3 mM KH_2PO_4, and 55 mM glucose, pH 8.0), centrifuged at 1600 *g* for 10 min at +4°C, and resuspended in Dignam Buffer A (Dignam *et al.*, 1983). The procedure to prepare S100 trypanosome extract is identical to that for HeLa cells (Dignam *et al.*, 1983). Extracts were stored at −70°C.

Cell growth and RNA extraction

Culture-form or procyclic trypanosomes were grown at 27°C without CO_2 in SDM-79 medium (Brun and Schonenberger, 1979) supplemented with 10% fetal calf serum, hemin, antimycotics and antibiotics. Cells were split 1:20 every 3−5 days. The doubling time is ~18 h, and at stationary phase the cells reach a density of 2.5×10^7/ml. For the phosphate labeling, ~100 ml of cells at a density of 1×10^7/ml were centrifuged at 1600 *g* for 10 min, and then resuspended in 20−25 ml of low phosphate (i.e. reduced 20-fold) SDM-79 medium containing 25−30 mCi of $[^{32}P]H_3PO_4$ (carrier-free, from ICN Radiochemicals). Trypanosomes were labeled for 4 h at 27°C without CO_2. Cells were harvested by centrifugation at 1600 *g* for 10 min and resuspended in either water/phenol (1:1) and frozen at −70°C or resuspended in 500 μl of 150 mM NaCl, 1% SDS, 100 mM Tris−HCl (pH 7.5) and 12.5 mM EDTA plus 1 mg of proteinase K, incubated at 37°C for 15 min, phenol/chloroform (1:1) extracted, ether extracted twice and ethanol precipitated. In the cases where cells were frozen in phenol:water, lysed cells were thawed, the aqueous phase was extracted three times with ether, and nucleic acid was precipitated with either 2.5 vol of ethanol or 1 vol of isopropanol. In either extraction procedure, the precipitations were carried out at −20°C overnight. RNA was separated into poly(A)$^-$ and poly(A)$^+$ fractions by batch hybridization to type III oligo(dT)−cellulose (Collaborative Research) as described by Nakazato and Edmonds (1974).

Chemical and enzymatic treatment of RNA

Labeled branch from trypanosomes was isolated by performing an RNase protection essentially as described (Myers *et al.*, 1985) on uniformly labeled poly(A)$^-$ RNA using the anti-sense RNA probe shown in Figure 1. Protected RNA was electrophoresed on a 7 M urea−5% polyacrylamide gel, and the species at ~75 nucleotides was electroeluted into a dialysis sac. The eluted RNA was dialyzed against a large volume of water for several hours at 4°C, and the water was evaporated under vacuum. The resultant RNA was then treated in a small volume with either 1 μg of nuclease P1 in 75 mM ammonium acetate pH 5.0 or with 100 U of RNase T1, 1 μg of RNase A, and 5 U RNase T2 in 20 mM sodium acetate pH 4.6 for 2−4 h at 37°C. Resultant species were separated on two-dimensional cellulose F TLC plates using concentrated isobutyric acid/NH_4OH/H_2O (585:19:396) in the first dimension and saturated ammonium sulfate/isopropanol/1 M sodium acetate (80:2:18) in the second dimension (Silberklang *et al.*, 1979).

Most digests of recovered nucleotides were performed for 2 h at 37°C in 5−10 μl. The nuclease P1 digests were performed in 75 mM ammonium acetate (pH 5.0) as described above. The alkaline phosphatase digestions were performed in water with 3−6 μg of enzyme. RNA species were 5′ end-labeled using polynucleotide kinase (US Biochemicals) and crude $[\gamma\text{-}^{32}P]$ATP (ICN Radiochemicals) under standard reaction conditions. The sodium periodate oxidation and β-elimination reactions were carried out essentially as described (Sninsky *et al.*, 1976; Wallace and Edmonds, 1983). Samples were desalted using TEAB prior to two-dimensional chromatography.

Lariat RNA was resuspended in water or TE (10 mM Tris−HCl, pH 8.0, 1 mM EDTA), and debranching reactions and RNA extractions were performed essentially as previously described (Ruskin and Green, 1985; Sutton and Boothroyd, 1986). In cases where the amount of S100 was < 10% of the reaction volume, the debranching buffer contained 1 mg/ml of RNase-free BSA (BRL). Markers for denaturing polyacrylamide gels were *Msp*I fragments of pAT153 ^{32}P end-labeled using DNA polymerase large fragment.

Acknowledgements

We thank Phillip Zamore and Dr Michael Green of Harvard University for the generous gift of lariat RNA and HeLa cell S100 extract, Dr Mary Edmonds of University of Pittgsburgh for the chemically synthesized tetraribonucleotide branch and Dr Rashid Aman for the initial culture of procyclic trypanosomes. This work was supported in part by grants from the National Institutes of Health (AI21023) and the MacArthur Foundation. R.E.S. was supported by the Medical Scientist Training Program of the NIH (GM07365) and by a training grant from the Johnson and Johnson Co. J.C.B. holds a Molecular Parasitology Award from the Burroughs-Wellcome Fund.

References

Arenas,J. and Hurwitz,J. (1987) *J. Biol. Chem.*, **262**, 4274−4279.
Boothroyd,J.C. (1985) *Annu. Rev. Microbiol.*, **39**, 475−502.
Boothroyd,J.C. and Cross,G.A.M. (1982) *Gene*, **20**, 279−287.
Borst,P. (1986) *Annu. Rev. Biochem.*, **55**, 701−732.
Brun,R. and Schonenberger,M. (1979) *Acta Trop.*, **36**, 289−292.
Campbell,D.A., Thornton,D.A. and Boothroyd,J.C. (1984) *Nature*, **311**, 350−355.
De Lange,T., Berkvens,T.M., Veerman,H.J.G., Carlos,A., Frasch,C., Barry,J.D. and Borst,P. (1984) *Nucleic Acids Res.*, **12**, 4431−4443.
De Lange,T., Liu,A.Y.C., Van der Ploeg,L.H.T., Borst,P., Tromp,M.C. and Van Boom,J.H. (1983) *Cell*, **34**, 891−900.
De Lange,T., Michels,P.A.M., Veerman,H.J.G., Cornelissen,A.W.C.A. and Borst,P. (1984) *Nucleic Acids Res.*, **12**, 3777−3789.
Dignam,J.D., Lebowitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475−1489.
Dorfman,D. and Donelson,J. (1984) *Nucleic Acids Res.*, **12**, 4907−4920.
Freistadt,M., Robertson,H.D., Branch,A. and Cross,G.A.M. (1987) *Nucleic Acids Res.*, **15**, 9861−9880.
Glass,D.J., Polvere,R.I. and Van der Ploeg,L.H.T. (1986) *Mol. Cell. Biol.*, **6**, 4657−4666.
Green,M.R. (1986) *Annu. Rev. Genet.*, **20**, 671−708.
Guyaux,M., Cornelissen,A.W.C.A., Pays,E., Steinert,M. and Borst,P.

(1985) *EMBO J.*, **4**, 995–998.
Hornig,H., Aebi,A. and Weissmann,C. (1986) *Nature*, **324**, 589–591.
Jacquier,A. and Rosbash,M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5835–5839.
Kierzek,R., Kopp,D.W., Edmonds,M. and Caruthers,M.H. (1986) *Nucleic Acids Res.*, **14**, 4751–4764.
Konarska,M.M., Grabowski,P.J., Padgett,R.A. and Sharp,P.A. (1985) *Nature*, **313**, 552–557.
Kooter,J., De Lange,T. and Borst,P. (1984) *EMBO J.*, **3**, 2387–2392.
Krause,M. and Hirsch,D. (1987) *Cell*, **49**, 753–761.
Laird,P.W., Zomerdijk,J.C.B.M., Korte,D.D.E. and Borst,P. (1987) *EMBO J.*, **6**, 1055–1062.
Langford,C.J., Klinz,F.-J., Donath,C. and Gallwitz,D. (1984) *Cell*, **36**, 645–653.
Lanham,S.M. and Godfrey,D.G. (1970) *Exp. Parasitol.*, **28**, 521–534.
Michiels,F., Matthyssens,G., Kronenberger,P., Pays,E., Dero,B., Van Assel,S., Darville,M. and Cravador,A. (1983) *EMBO J.*, **2**, 1185–1192.
Milhausen,M., Nelson,R.G., Sather,S., Selkirk,M. and Agabian,N. (1984) *Cell*, **38**, 721–729.
Muhich,M.L., Hughes,D.E., Simpson,A.M. and Simpson,L. (1987) *Nucleic Acids Res.*, **15**, 3141–3153.
Murphy,W.J., Watkins,K.P. and Agabian,N. (1986) *Cell*, **47**, 517–525.
Myers,R.M., Larin,Z. and Maniatis,T. (1985) *Science*, **230**, 1242–1246.
Nakazato,H. and Edmonds,M. (1974) *Methods Enzymol.*, **29**, 431–443.
Nelson,R.G., Parsons,M., Barr,P.J., Stuart,K., Selkirk,M. and Agabian,N. (1983) *Cell*, **34**, 901–909.
Nelson,R.G., Parsons,M., Selkirk,M., Newport,G., Barr,P.J. and Agabian,N. (1984) *Nature*, **308**, 665–667.
Newman,A.J., Lin,R.-J., Cheng,S.-C. and Abelson,J. (1985) *Cell*, **42**, 335–344.
Padgett,R.A., Grabowski,P.J., Konarska,M.M., Seiler,S. and Sharp,P.A. (1986) *Annu. Rev. Biochem.*, **55**, 1119–1150.
Parsons,M., Nelson,R.G., Watkins,J.B. and Agabian,N. (1984) *Cell*, **38**, 309–316.
Peebles,C.L., Perlman,P.S., Mecklenburg,K.L., Petrillo,M.L., Tabor,J.H., Jarrel,K.A. and Cheng,H.-L. (1986) *Cell*, **44**, 213–223.
Perry,K.L., Watkins,K.P. and Agabian,N. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8190–8194.
Ruskin,B. and Green,M.R. (1985) *Science*, **229**, 135–140.
Silberklang,M., Gillum,A.M. and Rajbhandary,U.L. (1979) *Methods Enzymol.*, **59**, 58–109.
Sninsky,J.J., Last,J.A. and Gilham,P.T. (1976) *Nucleic Acids Res.*, **3**, 3157–3167.
Sogin,M.L., Elwood,H.J. and Gunderson,J.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1383–1387.
Sutton,R.E. and Boothroyd,J.C. (1986) *Cell*, **47**, 527–535.
Sutton,R.E. and Boothroyd,J.C. (1988) *Mol. Cell. Biol.*, **8**, 494–496.
Van der Ploeg,L.H.T., Cornelissen,A.W.C.A., Michels,P.A.M. and Borst,P. (1984) *Cell*, **39**, 213–221.
Van der Ploeg,L.H.T., Liu,A.Y.C., Michels,P.A.M., De Lange,T., Borst,P., Majumder,K., Weber,H. and Veeneman,G.H. (1982) *Nucleic Acids Res.*, **10**, 3591–3604.
Van der Veen,R., Arnberg,A.C., Horst,G., Bonen,L., Tabak,H.F. and Grivell,L.A. (1986) *Cell*, **44**, 225–234.
Wallace,J.C. and Edmonds,M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 950–954.
Yost,H.J. and Lindquist,S. (1986) *Cell*, **45**, 185–193.

Received on December 29, 1987; revised on February 2, 1988