MOLECULAR AND CELLULAR BIOLOGY, Jan. 1988, p. 494–496
0270-7306/88/010494-03$02.00/0

Vol. 8, No. 1

The Cap of Both Miniexon-Derived RNA and mRNA of Trypanosomes Is 7-Methylguanosine

RICHARD E. SUTTON AND JOHN C. BOOTHROYD*

Department of Medical Microbiology, Stanford University School of Medicine, Stanford, California 94305

Received 21 August 1987/Accepted 14 October 1987

Most, if not all, trypanosome mRNAs have the same 35-base sequence at their 5′ terminus which is derived from a short RNA (medRNA) probably by the process of *trans*-splicing. It is of interest, evolutionarily and mechanistically, to determine the chemical structure of the 5′ terminus of the precursor (medRNA) and product (mRNA). We demonstrate here that the cap structure of both is most probably 7-methylguanosine in a 5′,5′ triphosphate linkage, consistent with a precursor/product relationship.

Most, if not all, mature mRNAs of *Trypanosoma brucei* have an identical 35-nucleotide miniexon or spliced-leader sequence at the 5′ terminus (2, 4). This sequence is derived from an RNA species of ~137 nucleotides, the first 35 of which correspond to the miniexon (7, 10, 15). This short RNA (named miniexon-derived or medRNA) is transcribed from a genomic DNA sequence of 1.35 kilobases which is tandemly repeatedly ~200 times in the trypanosome genome (9, 21). We and others have recently presented evidence that the miniexon is *trans*-spliced onto the protein-coding exons of acceptor RNAs to yield mature mRNAs (13, 18, 25). An intermediate in this process is a Y structure, composed of the tail portion of medRNA covalently attached to the upstream intron of the acceptor RNA. *trans*-splicing is probably a routine physiological process in all kinetoplastid parasites (8, 16, 22), and it may also occur to a limited extent in higher eucaryotes (e.g., *Caenorhabditis elegans*) (11).

Because Kinetoplastida occupy a unique branch of the evolutionary tree and the structure of medRNA probably plays an important role in the process of *trans*-splicing, we were interested in determining the chemical identity of the cap of both medRNA and mature mRNA molecules. Previously, it had been reported that medRNA is capped (12, 14). Here we show that the cap of both medRNA and mRNA of trypanosomes is 7-methylguanosine (7-meG) in a 5′,5′ triphosphate linkage.

Procyclic (insect-stage) trypanosomes were grown at 27°C without CO_2 in SDM-79 medium (5) supplemented with 10% fetal calf serum. For phosphate labeling, roughly 100 ml of late-log cells at a density of 10^7/ml was centrifuged at 1,600 × *g* for 10 min, suspended in 20 to 25 ml of low-phosphate SDM-79 medium (5) containing 25 to 30 mCi of [^{32}P]H_3PO_4 (carrier-free, from ICN Pharmaceuticals Inc.), and labeled for 4 h at 27°C without CO_2. Cells were harvested, and total RNA was prepared by extraction with phenol and precipitation with isopropanol. RNA was separated into poly(A)$^+$ and poly(A)$^-$ fractions by oligo(dT) batch chromatography as described previously (20).

The cap of medRNA was prepared from poly(A)$^-$ RNA by RNase protection with a probe which spans the start site of transcription of medRNA (−275 to +110 relative to the start site [Fig. 1a]). The 110-nucleotide protected portion of medRNA was electroeluted from a denaturing polyacrylamide gel (Fig. 1b), dialyzed against water, and dried. It was then treated with RNase A (Sigma Chemical Co.), T1 (Bethesda Research Laboratories, Inc.), and T2 (Sigma) (19) and chromatographed on cellulose F thin-layer chromatographic (TLC) plates (24) with 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 (Fig. 1c). Only five major spots were

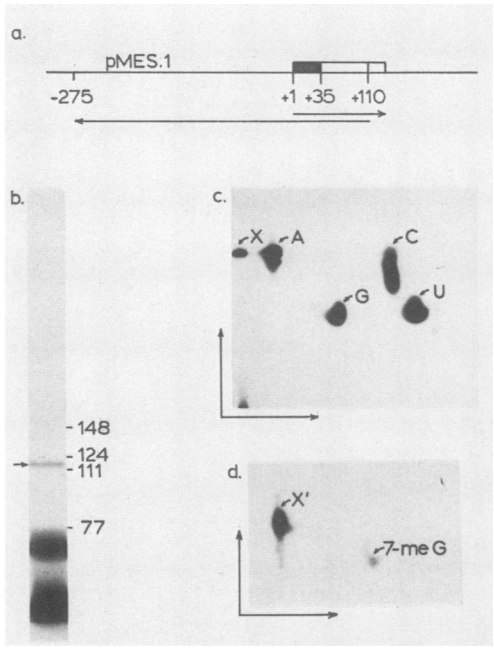

FIG. 1. (a) Schematic map of a cloned miniexon repeat (pMES.1 [7]) showing the miniexon (▬), miniintron (▭), medRNA transcript (→), and the protecting antisense RNA (←). Nucleotides are numbered from the published 5′ end of medRNA (7). (b) 110-nucleotide product of RNase protection (indicated by arrow) resolved by denaturing polyacrylamide gel electrophoresis. Markers (in nucleotides) are *Msp*I fragments of pAT153. (c) Two-dimensional cellulose F TLC of the RNase A, T1, and T2 digest of the 110-nucleotide fragment of medRNA (Fig. 1b). The four 3′-nucleoside monophosphates are indicated; X is the RNase-resistant cap. The first dimension is vertical, and the second dimension is horizontal. (d) Two-dimensional cellulose F TLC of the TAP-treated core cap (Fig. 1c). The released 7-meGMP is indicated as 7-meG; X′ is the remainder of the core cap less one phosphate (not visible in this exposure). The first dimension is vertical, and the second dimension is horizontal.

* Corresponding author.

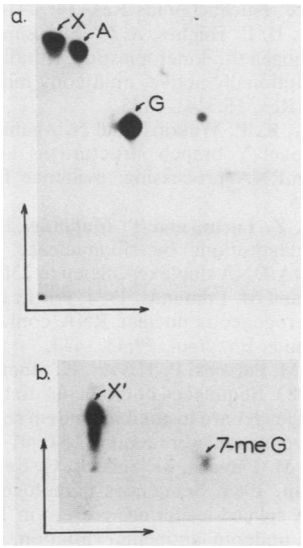

FIG. 2. (a) Two-dimensional cellulose F TLC of core cap eluted from the boronate-polyacrylamide column. (b) Two-dimensional cellulose F TLC of TAP-treated core cap eluted from the TLC plate (Fig. 2a). Chromatographed untreated core cap showed no spots other than the eluted starting material (data not shown). Other details are as described in the legend to Fig. 1.

reproducibly observed, four of which correspond to the 3'-nucleoside monophosphates. The fifth spot had the mobility expected of an RNase-resistant cap structure.

To analyze the cap of mature mRNA, poly(A)$^+$ RNA was digested with RNase A, T1, and T2 and chromatographed on a boronate-polyacrylamide column (Pierce Chemical Co.) as previously described (23). Capped oligonucleotides were eluted by applying a solution of 200 mM NaCl and 50 mM sodium acetate (pH 5.0), desalted with triethylamine bicarbonate, and chromatographed as described above on a two-dimensional cellulose F TLC plate (Fig. 2a). Only three major radioactive spots were observed, two of which correspond to 3'-AMP and -GMP; the third was the presumptive RNase-resistant cap of mRNA.

The putative RNase-resistant caps of medRNA and mRNA were scraped from the plates, eluted, dried, treated with tobacco acid pyrophosphatase (TAP; Promega Biotec), and chromatographed on cellulose F TLC plates (Fig. 1d and 2b). TAP only cleaves eucaryotic caps between adjacent phosphates, releasing 5'-nucleoside monophosphate and free phosphate. In both cases, a radioactive species which comigrates with a 5' 7-meGMP standard was released, in addition to a small amount of free P_i (not evident in these exposures because it was distributed over a large area). The coincident migration in this two-dimensional system with a 5' 7-meGMP standard run on the same plate strongly suggests identity. To our knowledge, under these chromatographic conditions, no other modified nucleotides have this precise mobility (24). In particular, it should be noted that 5' 2,2,7-trimethylguanosine monophosphate, the cap found on most eucaryotic small nuclear RNAs, has a dramatically different mobility, running much faster in the first dimension and much slower in the second (27). Together with the precedent of mRNA caps being 7-meG for most other eucaryotic mRNAs (1), the data presented here lead us to conclude that 7-meG is the true cap of trypanosome medRNA and mRNA.

Preliminary analyses using phosphodiesterase from the RNase-resistant species (Fig. 1c and 2a) indicate that it contained at least four other modified nucleotides, in addition to 7-meG (unpublished results). Three of these comigrated with 2'-*O*-methyladenosine, 2'-*O*-methylcytidine, and N^6-methyl, 2'-*O*-methyladenosine; the fourth did not comigrate with any standard in our possession. These results are consistent with the in vivo transcription start site of medRNA being either at the published position (3, 7) or four nucleotides upstream, as has recently been suggested (18).

We have presented evidence that the cap of both medRNA and mature mRNA of trypanosomes consists of 7-meG attached in a 5',5' triphosphate linkage. Based upon rRNA sequences, the Kinetoplastida separated from the main evolutionary tree before the split of fungi, plants, and animals (26). All genera in this order appear to share certain properties, including using a medRNA to *trans*-splice a miniexon sequence onto mRNAs (16, 17, 22). It will be of interest to see whether all the kinetoplastid medRNAs are capped by 7-meG. The data presented here indicate that the use of 7-meG as a cap is evolutionarily very ancient, and the fact that it is ubiquitous in nature implies an important function (for a review, see reference 1).

To our knowledge, medRNA is the shortest RNA molecule to be capped with 7-meG. These data suggest that capping with 7-meG need not be limited to large mRNA molecules. They further suggest medRNA may be functionally and evolutionarily unrelated to the U small nuclear RNAs of higher eucaryotes (6).

The fact that the caps of medRNA and mRNA are identical is consistent with the notion that medRNA serves as the cap donor for mRNAs. This may not be the sole function of the complex cap of medRNA; it may play important roles in secondary and tertiary structure formation of medRNA, in *trans*-splicing, and in translation of mRNAs. These questions will be difficult to address until DNA transformation or in vitro *trans*-splicing become available for trypanosomes.

Similar results to these have recently been obtained by others (K. L. Perry, K. P. Watkins, and N. Agabian, Proc. Natl. Acad. Sci. USA, in press; M. Freistadt, H. T. Robertson, A. Branch, and G. A. M. Cross, Nucleic Acids Res., in press). They further conclude that the start position for transcription is four nucleotides upstream from the published position.

We thank Mike Muhich, Robert Rhoads, and Norman Pace for useful comments and suggestions; Harold Drabkin, Rashid Aman, and Tim Mietzner for supplying materials; and Karen Perry and Marion Freistadt for exchange of information before publication.

This work was supported in part by Public Health Service grant AI21023 from the National Institutes of Health and grants from the MacArthur Foundation. R.E.S. was supported by the Medical Scientist Training Program (GM07365) of the National Institutes of Health and by a training grant from the Johnson and Johnson Co. J.C.B. holds a Molecular Parasitology Award from the Burroughs-Wellcome Fund.

LITERATURE CITED

1. **Banerjee, A. K.** 1980. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. Microbiol. Rev. **44**:175–205.
2. **Boothroyd, J. C.** 1985. Antigenic variation in african trypanosomes. Annu. Rev. Microbiol. **39**:475–502.
3. **Boothroyd, J. C., and G. A. M. Cross.** 1982. Transcripts coding for different variant surface glycoproteins of Trypanosoma brucei have a short, identical exon at their 5'-end. Gene **20**:281–289.
4. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55**:701–732.

5. **Brun, R., and M. Schonenberger.** 1979. Cultivation and in vitro cloning of procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Acta Trop. **36**:289–292.
6. **Busch, H., R. Reddy, L. Rothblum, and Y. C. Choi.** 1982. SnRNAs, snRNPs, and RNA processing. Annu. Rev. Biochem. **51**:617–654.
7. **Campbell, D. A., D. A. Thornton, and J. C. Boothroyd.** 1984. Apparent discontinuous transcription of Trypanosoma brucei variant surface antigen genes. Nature (London) **311**:350–355.
8. **De Lange, T., T. M. Berkvens, H. J. G. Veerman, A. Carlos, C. Frasch, J. D. Barry, and P. Borst.** 1984. Comparison of the genes coding for the common 5′ terminal sequence of messenger RNAs in three trypanosome species. Nucleic Acids Res. **12**:4431–4443.
9. **De Lange, T., A. Y. C. Liu, L. H. T. Van der Ploeg, P. Borst, M. C. Tromp, and J. H. Van Boom.** 1983. Tandem repetition of the 5′ mini-exon of variant surface glycoprotein genes: a multiple promoter for VSG gene transcription? Cell **34**:891–900.
10. **Kooter, J., T. De Lange, and P. Borst.** 1984. Discontinuous synthesis of mRNA in trypanosomes. EMBO J. **3**:2387–2392.
11. **Krause, M., and D. Hirsch.** 1987. A trans-spliced leader sequence on actin mRNA in C. elegans. Cell **49**:753–761.
12. **Laird, P. W., J. M. Kooter, N. Loosbroek, and P. Borst.** 1985. Mature mRNAs of Trypanosoma brucei possess a 5′ cap acquired by discontinuous RNA synthesis. Nucleic Acids Res. **13**:4253–4266.
13. **Laird, P. W., J. C. B. M. Zomerdijk, D. deKorte, and P. Borst.** 1987. In vivo labelling of intermediates in the discontinuous synthesis of mRNAs in Trypanosoma brucei. EMBO J. **6**:1055–1062.
14. **Lenardo, M. J., D. M. Dorfman, and J. E. Donelson.** 1985. The spliced leader sequence of *Trypanosoma brucei* has a potential role as a cap donor structure. Mol. Cell. Biol. **5**:2487–2490.
15. **Milhausen, M., R. G. Nelson, S. Sather, M. Selkirk, and N. Agabian.** 1984. Identification of a small RNA containing the trypanosome spliced leader: a donor of shared 5′ sequences of trypanosomatid mRNAs? Cell **38**:721–729.
16. **Miller, S. I., S. M. Landfear, and D. F. Wirth.** 1986. Cloning and characterization of a Leishmania gene encoding an RNA spliced leader sequence. Nucleic Acids Res. **14**:7341–7360.
17. **Muhich, M. L., D. E. Hughes, A. M. Simpson, and L. Simpson.** 1987. The monogenetic kinetoplastid Crithidia fasciculata contains a transcriptionally active, multicopy mini-exon sequence. Nucleic Acids Res. **15**:3141–3153.
18. **Murphy, W. J., K. P. Watkins, and N. Agabian.** 1986. Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: evidence for trans splicing. Cell **47**:517–525.
19. **Myers, R. M., Z. Larin, and T. Maniatis.** 1985. Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. Science **230**:1242–1246.
20. **Nakazato, H., and M. Edmonds.** 1974. Purification of messenger RNA and heterogeneous nuclear RNA containing poly(A) sequences. Methods Enzymol. **29**:431–443.
21. **Nelson, R. G., M. Parsons, P. J. Barr, K. Stuart, M. Selkirk, and N. Agabian.** 1983. Sequences homologous to the variant antigen mRNA spliced leader are located in tandem repeats and variable orphons in Trypanosoma brucei. Cell **34**:901–909.
22. **Nelson, R. G., M. Parsons, M. Selkirk, G. Newport, P. J. Barr, and N. Agabian.** 1984. Sequences homologous to the variant antigen mRNA spliced leader are present in Trypanosomatidae which do not undergo antigenic variation. Nature (London) **308**:665–667.
23. **Pace, B., and N. R. Pace.** 1980. The chromatography of RNA and oligoribonucelotides on boronate-substituted agarose and acrylamide. Anal. Biochem. **107**:128–135.
24. **Silberklang, M., A. M. Gillum, and U. L. Rajbhandary.** 1979. Use of in vitro ^{32}P labeling in the sequence analysis of nonradioactive tRNAs. Methods Enzymol. **59**:58–109.
25. **Sutton, R. E., and J. C. Boothroyd.** 1986. Evidence for trans splicing in trypanosomes. Cell **47**:527–535.
26. **Walder, J. A., P. S. Eder, D. M. Engman, S. T. Brentano, R. Y. Walder, D. S. Knutzon, D. M. Dorfman, and J. E. Donelson.** 1986. The 35-nucleotide spliced leader sequence is common to all trypanosome messenger RNAs. Science **233**:569–571.
27. **Wise, J. A., D. Tollervey, D. Maloney, H. Swerdlow, E. J. Dunn, and C. Guthrie.** 1983. Yeast contains small nuclear RNAs encoded by single copy genes. Cell **35**:743–751.