
RNA splicing is required to make the messenger RNA for a variant surface antigen in trypanosomes

L.H.T. Van der Ploeg, A.Y.C. Liu, P.A.M. Michels, T. De Lange, P. Borst, H.K. Majumder^{*†}, H. Weber^{*}, G.H. Veeneman[†] and J. Van Boom[†]

Section for Medical Enzymology and Molecular Biology, Laboratory of Biochemistry, University of Amsterdam, Jan Swammerdam Institute, P.O. Box 60.000, 1005 GA Amsterdam, The Netherlands

Received 29 April 1982; Accepted 21 May 1982

ABSTRACT

The expression of the gene for variant surface glycoprotein (VSG) 118 in *Trypanosoma brucei* is activated by transposing a DNA segment containing the gene and 1-2 kb in front of it to an expression site elsewhere in the genome. By S₁ nuclease protection and RNA blotting experiments we show here the presence of several minor transcripts in trypanosomes synthesizing VSG 118, one of which covers the entire transposed segment. Comparison of the sequence of the 5' terminal segment of VSG 118 messenger RNA (mRNA), determined by primed reverse transcription, and the corresponding region of the 118 VSG gene, shows that the 5' terminal 34 nucleotides of the mRNA are not encoded in the 118 VSG gene contiguous with the remainder of the mRNA. We conclude that synthesis of a VSG mRNA involves splicing of a much longer primary transcript, which may start outside the transposed segment.

INTRODUCTION

Trypanosoma brucei and related unicellular flagellates employ antigenic variation to evade the antibodies of their vertebrate host. Antigenic variation is mediated by successively expressing genes for different variants of the major surface antigen, the Variant Surface Glycoprotein (VSG). We have previously shown that the expression of two of these genes in stock 427 of *T. brucei* is controlled by a gene duplication-transposition [1-6]. The importance of this mechanism for control of VSG gene expression has been confirmed by others for two additional *T. brucei* strains [8-11].

A trypanosome variant expressing VSG gene 117 or 118 contains two copies of that gene, the basic copy (BC) present in all variants and the expression-linked extra copy (ELC), which is the one transcribed [5]. Comparison of BCs and ELCs of VSG genes 117 and 118 has defined the properties of ELC genes and their surrounding sequences (see Fig. 1). The transposed segment starts 1-2 kb in front of the BC gene and ends at the

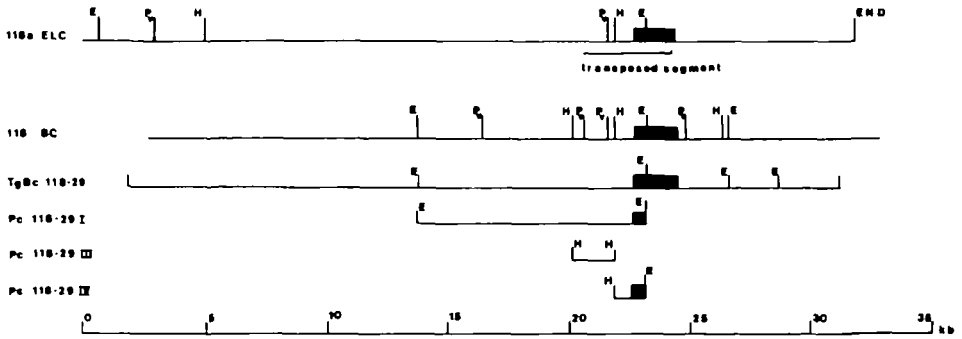


Fig. 1. Comparison of the VSG 118 BC and ELC physical maps with the 118 BC clones. Comparison of the restriction enzyme sites in the 118 BC and 118 ELC map shows the localization of the VSG 118 transposed segment. Below the 118 BC physical map the lengths of the 118 BC cosmid clone (TgBc 118-29) and the sub-clones derived from it are shown. A fragment containing the 10-kb EcoRI fragment with the 5'-end of the gene was inserted into pAT153. This clone (Pc 118-29-I) was then used to make the additional clones Pc 118-29-III and Pc 118 29-IV shown. See Materials and Methods for details. Abbreviations: E, EcoRI; H, HindIII and P_v, PvuII.

3'-end of the gene [6]; transposition is accompanied by a recombination which may lead to a replacement of the 3'-end of the duplicated gene [5-7]. In the expression site, the ELC is surrounded by long stretches of DNA that lack restriction enzyme sites. The size of these 'barren' regions is variable [7]. At the 3'-side the barren region ends at a discontinuity in the DNA, which could be the end of a chromosome (De Lange, T., unpublished). These unusual properties of the regions flanking ELCs may account for the failure of ELC clones to turn up in recombinant DNA clone banks from which clones containing BC genes were readily isolated [4-6].

In this paper we address the question how the duplication-transposition activates VSG gene expression. We have briefly reported (see refs 5,6,12) two observations indicating that the mature VSG mRNA is not the primary transcript of the ELC gene. First, we have observed minor transcripts that are variant-specific and that hybridize with DNA segments from the region in front of the expressed VSG gene. Second, we have found by S₁ nuclease protection experiments that the mRNA for VSG 118 protected a fragment in our longest complementary DNA (cDNA)

clone that was 10 bp longer at the 5'-side than the protected fragment derived from a 118 BC clone. This proved that the 5' terminus of the mRNA is not contiguously encoded with the rest of the mRNA sequence in the BC gene. This 5' terminus must, therefore, come from upstream and be joined to the body of the mRNA by RNA splicing.

In this paper we present our RNA blot experiments in detail and compare the sequence of the 5'-end of the mRNA and the BC gene of VSG 118. The results show that this VSG mRNA arises from a precursor RNA by a complex splicing process and suggests a mechanism for the control of VSG gene expression by transposition.

MATERIALS AND METHODS

RNA isolation - White male Wistar rats (300 g) were used to grow large batches of trypanosomes for RNA isolation. We have used the cloned variant antigen types MITat 1.4 (117), 1.5 (118a) and 1.186 (118c). The numbers in brackets designate the old names of these variants. These are used in the paper rather than the conventional nomenclature. Total trypanosomal RNA was isolated by LiCl precipitation, essentially as described by Auffray and Rougeon [13]. In short, rats were bled by cardiac puncture when a density of $1-3 \times 10^9$ trypanosomes/ml was reached. The total blood was immediately lysed in 10 volumes of 3 M LiCl, 6 M urea by strenuously mixing in a Sorvall omnimixer at 0°C. The RNA was allowed to aggregate at 0°C for 12 h and collected by centrifugation for 60 min at 20,000 rpm, 4°C in a Beckman Spinco SW27 rotor. The RNA was dissolved in 10 mM Tris-HCl (pH 7.5), 10 mM EDTA, 0.1% sodium dodecylsulphate and 10 µg proteinase K per ml, extracted with phenol and precipitated with ethanol. The RNA was dissolved in 10mM Tris-HCl (pH 7.5), 10 mM MgCl₂ and incubated for 15 min with pancreatic DNase I at 37°C. After DNase I treatment the RNA was again treated with proteinase K and phenol. The RNA was stored at -20°C as an ethanol precipitate.

Agarose gel electrophoresis, transfer of RNA to nitrocellulose filters and filter hybridization - Glyoxylated RNA samples (20-40 µg) were size-fractionated in 1% agarose gels at 0.15

V/cm for 1 h, followed by 1-3 V/cm for 6 h. The RNA was then transferred to nitrocellulose filters as described by Thomas [14]. The recombinant plasmids containing DNA complementary to VSG mRNA have been described by Hoeljmakers et al. [1]. The recombinant cosmid TgBc 118-29, containing the 118 BC gene (Van der Ploeg, L.H.T., unpublished), was used to make sub-clones of T. brucei DNA in the Escherichia coli plasmid pAT 153. One of these sub-clones contains a 1700-bp HindIII fragment, the other a 1410-bp HindIII-EcoRI fragment. Fig. 1 shows the location of these fragments in the map of the 118 BC gene. Sub-probes of cloned DNA were isolated by preparative agarose gel electrophoresis of the appropriate restriction endonuclease digests, followed by electrophoresis of the DNA into Whatman 3-MM paper and elution of the DNA from the paper as described [15]. DNA fragments used for sequence analysis were isolated from polyacrylamide gels according to Smith and Calvo [16].

The fragments used as radioactive probes were labelled by nick-translation [17]. Hybridization of filters at 65°C in 3 x SSC, 0.1% sodium dodecylsulphate, 50 µg single-stranded salmon-sperm DNA per ml, 10% dextran sulphate, 10 x Denhardt solution and 10 µg poly(A) per ml was as described previously [5].

Post-hybridizational washes at 65°C to remove aspecifically-bound probe and autoradiography were done as described [5].

S₁ nuclease protection experiments - The method of Berk and Sharp [18] was used, as described in Bernards et al. [5].

DNA sequence analysis - The sequence analysis of the 5'-end of the 118 BC VSG gene and the adjacent co-transposed segment was done on a sub-fragment of the HindIII-EcoRI sub-clone (see Fig. 1). DNA sequence analysis was performed using the chemical degradation procedures as described by Maxam and Gilbert [19], with the purine reaction as described by Cooke et al. [20].

RNA sequence analysis - Poly(A) RNA was isolated as outlined by Bernards et al. [5]. Primed cDNA synthesis, using a synthetic dodecadeoxynucleotide as primer, with chain terminators was carried out essentially as described [21,22]. The ratio of ddATP and ddCTP to dATP and dCTP, respectively, was one to one and that of ddGTP and ddTTP to dGTP and dTTP was four to one. The three non-radioactive deoxynucleotide triphosphates were at a concentration of 100 µM. The labelled triphosphate was at a

concentration of 2.5 μM (400 Ci/mmol, The Radiochemical Center, Amersham). About 1 unit of reverse transcriptase was used in each reaction which contained 0.1 μg VSG-specific mRNA and 10-40 ng primer. The reaction was carried out at 42 $^{\circ}\text{C}$ for 15 min and chased with 100 μM of the dNTPs for a further 15 min. The reaction was stopped by adding an equal volume of 80% formamide, 10 mM NaOH, 1 mM EDTA, 0.1% bromphenol blue and 0.1% xylene cyanol FF. The total volume of 10 μl was then heated to above 90 $^{\circ}\text{C}$ for 3-5 min before being loaded onto gel. Exposure was on Kodak X-1 film with intensifying screen for 12-48 h at -70 $^{\circ}\text{C}$.

It should be noted here that an error in the initial BC gene sequence misled us to construct a primer with a mismatched base at the 3'-end. In hybrid with the mRNA a CT pairing will result. Since there is no evidence for 3'-5' exonuclease activity of reverse transcriptase, it is likely that the mismatch was by-passed. The sequence obtained shows that priming was specific and started beyond the mismatch.

Synthesis of the oligonucleotide primer - The DNA fragment was prepared on the solid support polystyrene via a phosphotriester approach [23]. The crude fragment was purified by Sephadex G50 column chromatography and converted into the ammonium salt by passing it through a column containing Dowex 50W cation exchange resin (NH_4^+ form). Enzymatic digestion of the fragment with venom phosphodiesterase and High Performance Liquid Chromatography analysis [24] of the digest showed complete digestion and the presence of solely dT, pdC, pdA, pdG and pdT in the correct ratios.

RESULTS

RNA blot analysis of transcripts of the 118 VSG gene and its co-transposed flanking region

To minimize the possible loss of labile precursor RNAs, the RNA preparations used for hybridization analysis were isolated directly from trypanosomes in rat blood without taking the trypanosomes through the lengthy standard purification procedure. The RNA was glyoxylated and size-fractionated in agarose gels and transferred to nitrocellulose filters. The

hybridization of these RNA blots with various DNA probes is illustrated in Fig. 2. Probe 1, which covers the 5'-edge of the transposed segment, detects three major transcripts (4000, 1800 and 1600 nucleotides) and one minor transcript (1050 nucleotides) in the 118 RNA sample (lane B). Transcripts with the same mobility were also detected by probe 2, which covers the segment directly in front of the gene. In addition, this probe detects transcripts at 3800, 2900, 2400 and 600 nucleotides. Finally, the cDNA probe denoted 3 detects the transcripts of 4000, 3800, 2900 and 2400, as well as a 5000-nucleotide transcript not seen by the other probes, and the mature mRNA (2100 nucleotides).

All transcripts detected by these probes under stringent

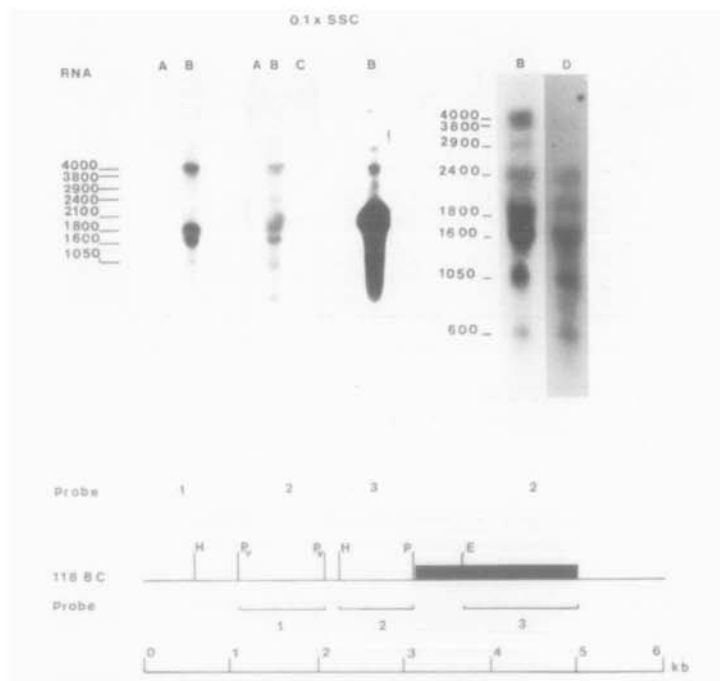


Fig. 2. Identification of minor transcripts of the VSG 118 gene by RNA blot analysis. 20-40 μ g of glyoxylated RNA was size-fractionated in 1% agarose gels and transferred to nitrocellulose filters. The RNA samples used are: A, 117a RNA; B, 118a RNA; C, *T. brucei* strain 427 RNA from *in vitro* cultured trypanosomes; D, 118c RNA. The 118 BC physical map below the figure shows the map location of the probes used for identification of transcripts.

conditions are absent in RNA from variant 117 (lane A), which makes VSG 117 instead of VSG 118; and in RNA from culture-form T. brucei (lane C; only shown for probe 2), which do not possess a surface coat. This shows that the transcripts arise in a variant-specific manner, in agreement with our previous observation that the variant-specific ELC and not the BC is transcribed [5].

On the basis of the RNA blots, only tentative map positions can be given for these transcripts. The 4000-nucleotide RNA is probably transcribed from the entire transposed segment (2.7-3.7 kb), because it is detected by all three probes. Since the 3800-nucleotide RNA hybridizes more strongly to probes 2 and 3 than to probe 1, it probably covers only part of the PvuII fragment that contains the 5'-edge of the transposed segment. The smaller transcripts cannot yet be localized precisely.

The right-hand panel of Fig. 2 compares the hybridization of probe 2 with RNAs from two trypanosome variants that express the same VSG gene. Variant 118a (marked B) is the one used in the other experiments, variant 118c (marked D) has been isolated more recently. It has an ELC map that resembles that of 118a, but the 118a and 118c ELCs differ in the size of the 'barren' flanking sequences, which lack restriction enzyme cleavage sites (Michels, P.A.M., unpublished). The 118c RNA contains a similar set of RNAs complementary to probe 2 as the 118a RNA, but the relative amounts differ. The larger RNAs are present in relatively low amounts when compared to the 600-nucleotide transcript. This may result from minor differences in the stage of infection at which the trypanosomes are lysed or in differences introduced during the RNA isolation procedure.

Mapping of VSG gene 118 transcripts by S_1 nuclease protection experiments

To verify whether transcripts covering the entire co-transposed segment are present, appropriate DNA fragments were hybridized to RNA from variant 118 and the DNA segments protected against S_1 nuclease digestion were analysed on gels. Fig. 3 shows that the 1410-bp HindIII-EcoRI fragment, which extends from within the gene to 907 bp upstream, is completely protected by the 118 RNA. In addition, the blot shows the

550-bp fragment expected from the hybrid with mature 118 mRNA and two additional transcripts that must represent partial copies of the DNA fragment. The 850 fragment might arise from mixed hybrids that contain the 4000- or 3800-nucleotide transcripts and the 5'-end of the mature mRNA. It could also be part of a processing product derived from splicing of the VSG precursor RNA. The precise localization of the 1050-nucleotide transcript is unknown, but it could arise from the 2900-nucleotide or 2400-nucleotide transcripts detected with probes 2 and 3 in Fig. 2.

To localize the 5'-ends of the 4000-, 1800- and 1600-nucleotide transcripts detected in Fig. 2, S_1 nuclease experiments were done with DNA fragments covering the 5'-edge of the transposed segment. The middle panel of Fig. 3 shows that the HindIII fragment gives rise to 560- and 490-nucleotide transcripts. These are both shortened by 175 nucleotides in the right-hand panel where the PvuII fragment is used in the hybridization. 175 nucleotides is exactly the distance between the PvuII and HindIII cleavage sites in the transposed segment. This locates the two transcripts as indicated in the lower half of Fig. 3.

The main conclusion from these experiments is that there are transcripts, specific for variant 118, that cover most of the transposed segment. The results are compatible with the 4000-nucleotide transcript being a precursor RNA that is processed in a complex fashion to yield the other transcripts. The exact relation of the 4000-, 1800- and 1600-nucleotide transcripts (shown in Fig. 2) to the 560- and 490-nucleotide protected fragments (shown in Fig. 3) is, however, uncertain. The 5' terminal sequence of VSG 118 mRNA is spliced onto the remainder of the mRNA

Sequences at the 5'-end of the 118 BC gene and the cDNA plasmid TcV 118-2 were determined by sequence analysis of appropriate DNA restriction fragments (Fig. 4). This showed that the two sequences are identical but for the most 5' 9 bp of the cDNA, which are not contiguously encoded with the remainder of the sequence in the BC gene, confirming the S_1 nuclease protection experiments of Bernardis et al. [5]. To determine the size and sequence of the mRNA segment which is

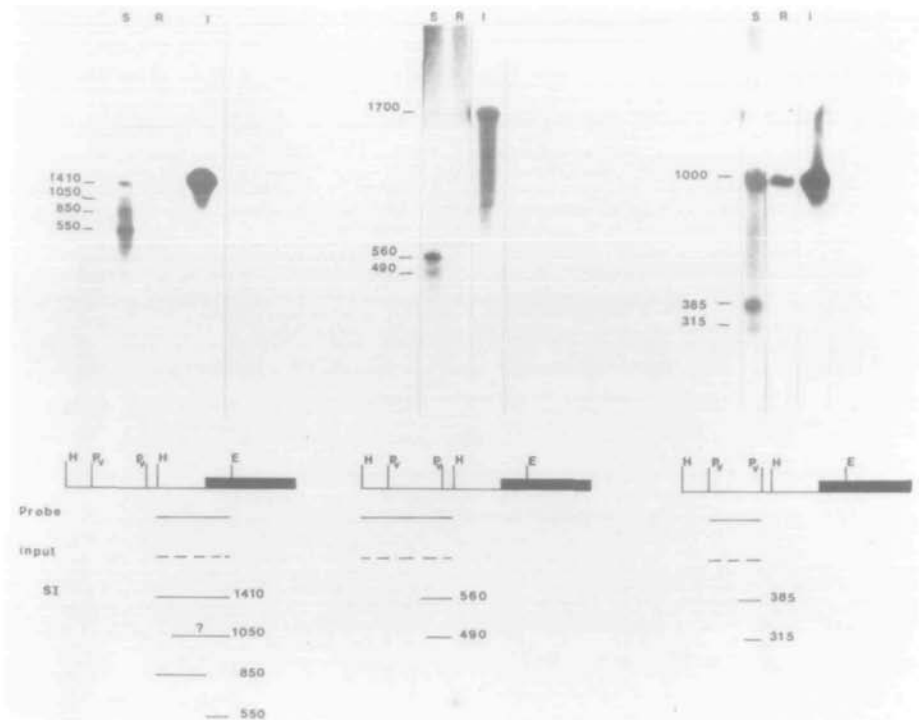


Fig. 3. S_1 nuclease protection analysis of minor transcripts of the VSG 118 gene. The experiments were done with 118a RNA and the DNA fragments and probes indicated below the physical maps were prepared as described under Methods. The length of the protected fragments (in nucleotides) and their location are outlined below the different panels. Abbreviations: S, S_1 nuclease treated; R, renaturation control; I, DNA input fragment. Other abbreviations are as in Fig. 1.

absent in the BC gene, the 5' terminal sequence of the mRNA was analysed by the chain-termination method [25]. A synthetic primer of 12 deoxynucleotides was used to specifically prime DNA synthesis by reverse transcriptase with total poly(A)⁺ RNA as template. An autoradiogram of a 118 RNA sequence gel is shown in Fig. 5 and the deduced sequence in Fig. 4. There is an unambiguous strong stop at 45 nucleotides beyond the primer. The sequence continues specifically - though only faintly visible - beyond this stop. We attribute this to the presence of longer transcripts, presumably the precursor RNAs identified in the preceding sections. The most 5' 34 nucleotides up to the

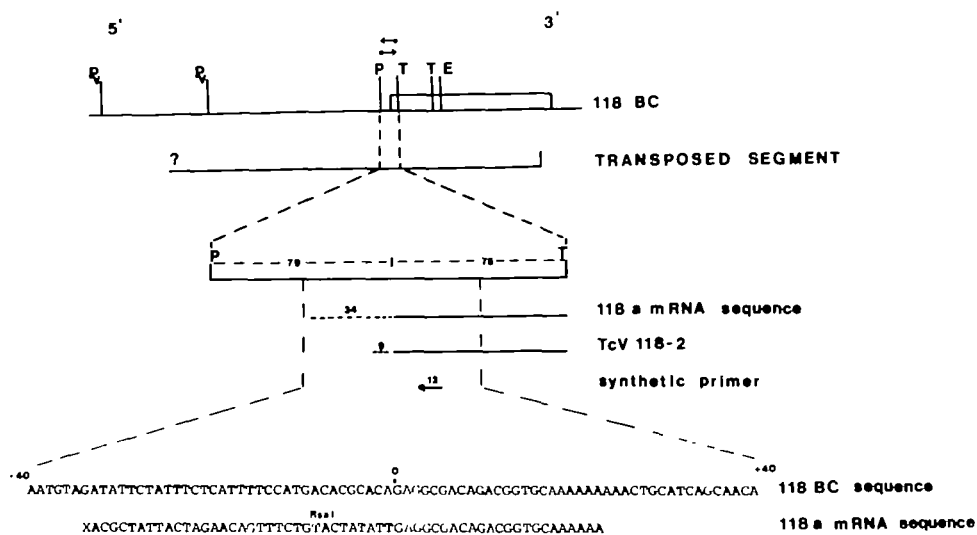


Fig. 4. Sequence comparison of the 118 BC and mRNA at the 5' terminus. The strategy for determining the BC sequence extending from the PstI and TaqI sites is denoted by the arrows above the schematic drawing of the 118 BC gene. The sequence determined from primed reverse transcription of the mRNA contains two uncertainties (see Fig. 5). X denotes the strong stop that appeared in every lane of the sequencing run. No other preceding strong stops due to secondary structure were found. The number of nucleotides in this terminal segment is arbitrarily set at 34 because of the uncertainties mentioned.

strong stop differ from the corresponding region in the BC gene. We conclude from this that mature VSG 118 mRNA results from the splicing of a precursor RNA. Indications that this precursor starts beyond the 5'-edge of the transposed segment will be presented in the Discussion.

The sequence at the position where the BC gene starts to diverge from the mRNA sequence is 5' GCACA/G (Fig. 4). This differs from the consensus sequence found at the intron-exon (acceptor) border of eukaryotic introns (Py)_n(Py)AG/G [26]. Trypanosomes may, therefore, not use this sequence in splicing precursor RNAs. It is also possible, however, that the intron is surrounded by a direct repeat and that the splice point is downstream of the actual divergence point in Fig. 4. A single base shift would result in a splice acceptor site CACAG/A, which resembles the consensus sequence.

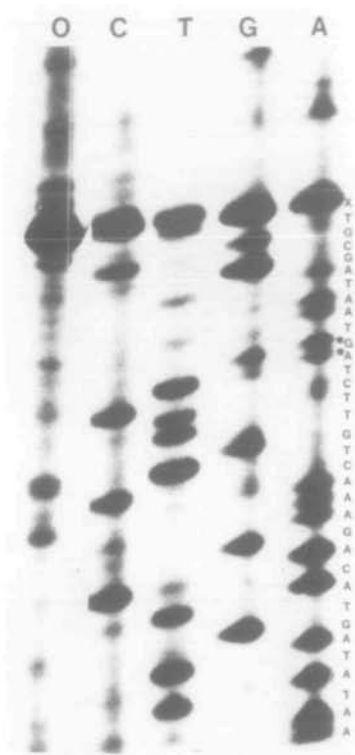


Fig. 5. The sequence at the 5' terminus of 118 mRNA starting at the point of divergence with respect to the BC sequence. The products of the primed reverse transcription were separated on an 8% polyacrylamide gel containing 7 M urea. The reaction mixture run in the O lane contained no dideoxy chain terminator. The sequence is complementary to the mRNA. There are two uncertainties as marked by the asterisks. X denotes the strong stop observed in all lanes. Extension products beyond this strong stop are also visible. They may be generated from priming on precursor molecules.

DISCUSSION

Our results show that the synthesis of VSG 118 mRNA must involve a longer precursor, which is spliced to yield a mature mRNA containing sequences that are non-contiguous in the DNA. Direct evidence for this conclusion comes from the sequence comparison of mRNA and BC gene, which shows that the first 34 nucleotides of the mRNA are not present in the BC gene. The presence of several transcripts longer and shorter than mature VSG mRNA and complementary to the region in front of the 118 gene is compatible with a complex splicing pathway with several intermediates and side products. These RNAs are all variant-specific and, therefore, linked to the expression of the VSG 118 gene. Whether they are all intermediates/side products of the main pathway for VSG mRNA synthesis remains to be determined. Since we have observed analogous minor transcripts of the VSG 117 gene, when it is expressed, we expect our results

for the VSG 118 gene to apply to VSG genes in general.

Where does the 5' terminal sequence of the 118 mRNA come from? We have recently found that two other independently-isolated trypanosome variants that express the same 118 gene (118b and 118c) make VSG mRNAs with the same 5' terminus as 118a (Liu, A.Y.C. and Michels, P.A.M., unpublished). This excludes the unlikely possibility that the 5' terminus is made by random nucleotide addition and shows that it is transcribed from a template that remains unaltered while genes move into and out of the expression site. Two observations argue against this template being part of the transposed segment: first, 90% of this sequence that is co-transposed to the expression site has been determined and in this sequence the 34 nucleotide terminus is not present (unpublished results). Second, the remaining 10% does not contain an *RsaI* site whereas the 5' terminus does (see Fig.4). The simplest interpretation of the data is therefore that the 34 bp come from outside the transposed segment. This would imply that the VSG gene transposition activates transcription of the gene by placing it downstream of a strong promoter. The synthesis of a precursor that is spliced to yield the mature mRNA would allow retention of the 5'-triphosphate end (required for capping?) and allow flexibility in the permissible size and sequence of the co-transposed segment. Pays et al. [9] have found, in fact, that the segment co-transposed with the AnTat 1.1 gene is smaller than 0.6 kb, i.e. much smaller than the 1-2 kb co-transposed with the 117 and 118 genes (Fig. 1).

Although splicing of precursors of rRNAs has been demonstrated in the protozoa *Tetrahymena* [27,28] and *Physarum* [29], this is to our knowledge the first example of the involvement of splicing in mRNA synthesis in protozoa.

The mechanism that we favour for the control of VSG gene expression is fundamentally different from that of the mating-type genes in yeast, to which VSG genes are often compared [1,4,12,30]. In the yeast system there is also an expression site to which duplicates of silent genes (cassettes) are transposed, but in that case gene transcription starts in the middle of the transposed segment and activation of transcription is

apparently due to a 'position' effect [31,32].

ACKNOWLEDGEMENTS

We thank Dr. R.Benne and Mr. A.Bernards for helpful comments on the manuscript and Dr. J.C.Boothroyd (The Wellcome Research Labs., Department of Immunochemistry, Beckenham, Kent BR3 3BS, U.K.) for an enlightening exchange of information on 5'-ends of VSG mRNAs. A.Y.C.L. received a fellowship from the Damon Runyon-Walter Winchell Cancer Fund (DRG-504). This work was supported in part by a grant to P.B. from The Foundation for Fundamental Biological Research (BION), which is subsidized by The Netherlands Organization for the Advancement of Pure Research (ZWO) and a grant from the World Health Organization (Nr. T16/181/T7/34).

Abbreviations: AnTat, Antwerp Trypanosome antigen type; BC, basic copy; bp, base pair(s); cDNA, complementary DNA; ELC, expression-linked copy; kb, kilo-base pair(s); MITat, Molteno Institute Trypanosome antigen type; mRNA, messenger RNA; rRNA, ribosomal RNA; SSC, 0.15 M NaCl, 0.015 M sodium citrate (pH 7.0); VSG, variant surface glycoprotein.

*Institut für Molekularbiologie, Universität Zürich, Honggerberg, 8093 Zürich, Switzerland.

†Organic Chemistry Laboratory, State University Leiden, Gorlaeus Laboratories, P.O. Box 9502, 2300 RA Leiden, The Netherlands.

‡Present address: Indian Institute of Chemical Biology, Calcutta 700032, India.

REFERENCES

- 1 Hoeljmackers, J.H.J., Borst, P., Van den Burg, J., Weissmann, C. and Cross, G.A.M. (1980) *Gene* 8, 391-417.
- 2 Borst, P., Frasnch, A.C.C., Bernards, A., Hoeljmackers, J.H.J., Van der Ploeg, L.H.T. and Cross, G.A.M. (1980) *Am.J.Trop.Med.Hyg.*, Suppl. 29, 1033-1036.
- 3 Frasnch, A.C.C., Bernards, A., Van der Ploeg, L.H.T., Borst, P., Hoeljmackers, J.H.J., Van den Burg, J. and Cross, G.A.M. (1980) In: *The Biochemistry of Parasites and Host-Parasite Relationships: The Host-Invader Interplay* (Van den Bossche, H., Ed.), North-Holland, Amsterdam, pp. 235-239.
- 4 Borst, P., Frasnch, A.C.C., Bernards, A., Van der Ploeg, L.H.T., Hoeljmackers, J.H.J., Arnberg, A.C. and Cross, G.A.M. (1981) *Cold Spring Harbor Symp.Quant.Biol.* 45, 935-943.
- 5 Bernards, A., Van der Ploeg, L.H.T., Frasnch, A.C.C., Borst, P., Boothroyd, J.C., Coleman, S. and Cross, G.A.M. (1981) *Cell* 27, 497-505.
- 6 Van der Ploeg, L.H.T., Bernards, A., Rijsewijk, F.A.M. and Borst, P. (1982) *Nucl.Acids Res.* 10, 593-609.
- 7 Michels, P.A.M., Bernards, A., Van der Ploeg, L.H.T. and Borst, P. (1982) *Nucl.Acids Res.* 10, in press.
- 8 Pays, E., Van Meirvenne, N., Le Ray, D. and Steinert, M. (1981) *Proc.Natl.Acad.Sci.U.S.* 78, 2673-2677.
- 9 Pays, E., Lheureux, M. and Steinert, M. (1981) *Nucl.Acids Res.* 9, 4225-4238.
- 10 Pays, E., Lheureux, M. and Steinert, M. (1981) *Nature* 292, 265-267.
- 11 Majiwa, P.A.O., Young, J.R., Donelson, J.E., Englund, P.T., Rovis, L. and Williams, R.O. (1982) *Nature*, submitted.

- 12 Borst, P. and Cross, G.A.M. (1982) *Cell*, in press.
- 13 Auffray, C. and Rougeon, F. (1980) *Europ.J.Biochem.* 107, 303-314.
- 14 Thomas, P.S. (1980) *Proc.Natl.Acad.Sci.U.S.* 77, 5201-5205.
- 15 Girvitz, S.C., Bacchetti, S., Rainbow, A.J. and Graham, F.L. (1980) *Anal.Biochem.* 106, 492-496.
- 16 Smith, D.R. and Calvo, J.M. (1980) *Nucl.Acids Res.* 8, 2255-2274.
- 17 Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) *J.Mol. Biol.* 113, 237-251.
- 18 Berk, A.J. and Sharp, P.A. (1977) *Cell* 12, 721-732.
- 19 Maxam, A.M. and Gilbert, W. (1980) In: *Methods in Enzymology* (Grossman, L. and Moldave, K., Eds), Vol. 65, Part I, Academic Press, New York, pp. 499-560.
- 20 Cooke, N.E., Coit, D., Weiner, R.I., Baxter, J.D. and Martial, J.A. (1980) *J.Biol.Chem.* 255, 6502-6510.
- 21 McGeoch, D.J. and Turnbull, N.T. (1978) *Nucl.Acids Res.* 5, 4007-4024.
- 22 Levy, S., Sures, I. and Kedes, L.H. (1979) *Nature* 279, 737-739.
- 23 Van der Marel, G.A., Wille, G., Marugg, J.E., De Vroom, E., Tromp, M., Van Boeckel, C.A.A. and Van Boom, J.H. (1982) *Rec.Trav.Chim.Pays-Bas*, in press.
- 24 De Rooij, J.F.M., Wille, G., Van Deursen, P.H., Serdijn, J. and Van Boom, J.H. (1979) *Rec.Trav.Chim.Pays-Bas* 98, 537-544.
- 25 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc.Natl.Acad. Sci.U.S.* 74, 5463-5467.
- 26 Mount, S.M. (1982) *Nucl.Acids Res.* 10, 459-472.
- 27 Wild, M.A. and Gall, J.G. (1979) *Cell* 16, 565-573.
- 28 Din, N., Engberg, J., Kaffenberger, W. and Eckert, W.A. (1979) *Cell* 18, 525-532.
- 29 Gubler, U., Wyler, T., Seebeck, T. and Braun, R. (1980) *Nucl. Acids Res.* 8, 2647-2664.
- 30 Strathern, J.N. and Herskowitz, I. (1979) *Cell* 17, 371-381.
- 31 Nasmyth, K.A., Tatchell, K., Hall, B.D., Astell, C. and Smith, M. (1981) *Nature* 289, 244-250.
- 32 Klar, A.J.S., Strathern, J.N., Broach, J.R. and Hicks, J.B. (1981) *Nature* 289, 239-244.