**The EMBO Journal** vol.4 no.4 pp.995–998, 1985

# *Trypanosoma brucei*: a surface antigen mRNA is discontinuously transcribed from two distinct chromosomes

**M.Guyaux, A.W.C.A.Cornelissen[1], E.Pays, M.Steinert and P.Borst[1]**

Université Libre de Bruxelles, Département de Biologie Moléculaires, Rue des Chevaux, 67, 1640 Rhode-St-Genèse, Belgium, and [1]The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis), Department of Molecular Biology, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

Communicated by M.Steinert

**The mRNAs for variant surface glycoproteins (VSGs) and many other proteins in *Trypanosoma brucei* start with the same sequence of 35 nucleotides, encoded by a separate mini-exon. There are ~200 mini-exon genes per trypanosome and these are highly clustered on large chromosomes. We have found two trypanosome variants that express a VSG gene located on a small, 225-kb chromosome. Each gene yields a mRNA containing the 35-nucleotide sequence even though the 225-kb chromosome does not contain a complete mini-exon gene. These results provide a strong support for the hypothesis that transcription of protein-coding genes in trypanosomes is discontinuous.**

*Key words:* chromosome/discontinuous transcription/surface antigen/telomere/trypanosome

## Introduction

In their mammalian host, African trypanosomes are covered by a surface coat, which consists of a single glycoprotein species, the variant surface glycoprotein (VSG). The trypanosome can change the antigenic composition of its surface by switching the VSG gene expressed (see Borst and Cross, 1982). Switching occurs either by a duplicative transposition of an inactive VSG gene to a telomeric expression site (Hoeijmakers *et al.*, 1980; Pays *et al.*, 1981, 1983a; Bernards *et al.*, 1981; Majiwa *et al.*, 1982; Longacre *et al.*, 1983), or by the activation of a new telomeric site and the inactivation of the old one (Laurent *et al.*, 1984; Pays *et al.*, 1983b; Bernards *et al.*, 1984; Michels *et al.*, 1984; Myler *et al.*, 1984).

The VSG genes are split in an unusual fashion: most of the mRNA is specified by single major exon; the first 35 nucleotides, however, are encoded in a so-called mini-exon (Van der Ploeg *et al.*, 1982; Boothroyd and Cross, 1982; De Lange *et al.*, 1983; Nelson *et al.*, 1983; Dorfman and Donelson, 1984). Although *Trypanosoma brucei* contains ~200 copies of this mini-exon, none of these is within 10 kb of any of the main exons of the VSG genes studied (De Lange *et al.*, 1983; Nelson *et al.*, 1983). This raised the question whether mini-exon and main exon are linked at all, a question underlined by the recent finding that many, if not all mRNAs in *T. brucei* start with the mini-exon sequence (De Lange *et al.*, 1984a; Parsons *et al.*, 1984a, 1984b).

Linkage of genes in *T. brucei* cannot be studied by classical genetics (although the genome is apparently diploid, the mechanism of genetic exchange is still an elusive process) or by chromosome cytology (no chromosome condensation). However, pulse field gradient gel electrophoresis (Schwartz *et al.*, 1983; Schwartz and Cantor, 1984), allows a partial fractionation of chromosome-sized DNA molecules (Van der Ploeg *et al.*, 1984a, 1984b, 1984c), as illustrated in Figure 1A. This technique uses alternating pulses of electricity in the North-South and in the West-East directions. To avoid shear degradation of large DNA, the trypanosomes are lysed and deproteinized in agarose blocks, which are placed in the gel slots. The DNA runs diagonally and separates into four fractions: (i) large DNA that remains close to the slot, (ii) about three chromosomes of 2000 kb (2 Mb), (iii) a series of small chromosomes between 200 and 2000 kb, (iv) ~100 mini-chromosomes of 50–150 kb.

Van der Ploeg *et al.* (1984a, 1984b) have previously shown that mini-exons are mainly located on large DNA in these gels. A small and variable fraction of mini-exons is found on 2-Mb chromosomes but none appear to be present in smaller chromosomes.

In contrast, telomeric VSG genes are mainly in the smaller chromosomes. This makes it possible to test whether such VSG genes, located on small chromosomes that lack the mini-exon, give rise to mRNAs with the mini-exon sequence. Such a gene was recently found by Van der Ploeg *et al.* (1984b), but for technical reasons their experiments were inconclusive. We have therefore screened the variant antigen gene repertoire of another *T. brucei* strain and found two genes expressed on a 225-kb chromosome.

## Results and Discussion

Figure 1 displays the analysis of chromosome-sized DNA molecules for three antigenic variants of strain 1125 of *T. brucei*. In variant AnTat 1.1A the AnTat 1.1 gene is activated by duplicative transpositions (Pays *et al.*, 1983a) and the resulting expression linked extra copy (ELC) is located in a 225-kb chromosome, as shown in Figure 1B, lane 1. Elution of this 225-kb band from the gel and digestion of the DNA with *Pst*I, yields the 2-kb 1.1A ELC fragment identified previously (Pays *et al.*, 1983a) (Figure 2A). The five other members of the AnTat 1.1 family, recognized by this probe and present in different *Pst*I fragments in Figure 2A, are absent from the 225-kb chromosome and must all be located in the large DNA.

Variant AnTat 1.3A in Figure 1 expresses the 1.3 VSG gene, variant AnTat 1.6A the 1.6 gene. Panel C of Figure 1 shows the same blot as in panel B, but hybridized with a probe for the 1.3 gene. This probe lights up the 2-Mb band in all three variants and an additional 225-kb band in DNA from variants 1.3A and 1.6A. Variant 1.3A arose from variant 1.1A and we have previously shown that the 1.3 VSG gene in variant 1.3A is activated by duplicative transposition (Laurent *et al.*, 1983). The restriction maps indicated that the new 1.3A ELC had replaced the 1.1A ELC in the telomeric expression site. The results in Figures 1 and 2 agree with this interpretation. The 225-kb DNA was eluted from a chromosome-fractionation gel and digested with *Sph*I. Lane 4 of Figure 2B shows that the 225-kb chromosome contains only the 14.8-kb *Sph*I fragment. We have previously shown that this fragment is preferentially digested with DNase I in isolated nuclei (Laurent *et al.*, 1983) and therefore corre-

© IRL Press Limited, Oxford, England.

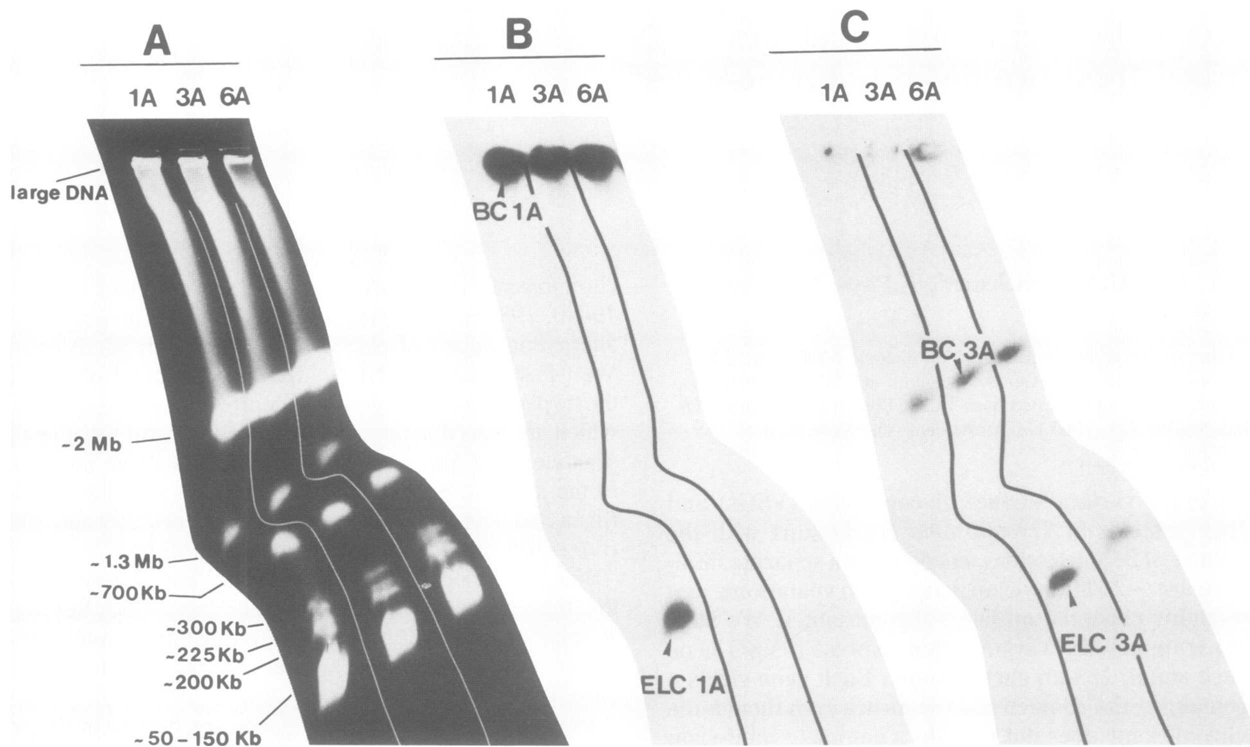


**Fig. 1.** Chromosomal location of the AnTat 1.1A and AnTat 1.3A VSG genes. Chromosome-sized DNA molecules of three different variant antigen types (indicated at the top of **panels A, B** and **C**) are separated using pulse field gradient (PFG) gel electrophoresis. The gel was stained with ethidium bromide (**panel A**) and then blotted onto nitrocellulose. The same filter was first hybridized with a $^{32}$P-labeled 1-kb *Pst*I insert of an AnTat 1.3A cDNA (Laurent *et al.*, 1983). The autoradiogram of the washed filter is shown in **panel C**. After removal of the probe with alkali, the filter was rehybridized with $^{32}$P-labeled 550-bp *Hin*dIII fragment of an AnTat 1.1A cDNA (Pays *et al.*, 1983a) giving the autoradiogram in **panel B**.

sponds to the ELC. This ELC is inactivated but not lost in the switch from variant 1.3A to 1.6A (Laurent *et al.*, 1984). Lane 3 of Figure 1C shows in addition that inactivation does not lead to a detectable change in chromosomal location.

The unambiguous localization of two successive ELCs on a small chromosome has allowed us to test whether a chromosome without mini-exon sequence can give rise to a mRNA that does contain this sequence. Figure 3 shows the outcome of this test. Nearly all hybridization with the mini-exon probe in panel B is to the large DNA, close to the slot. No specific hybridization is visible over the 225-kb chromosomes, whether an active ELC is present (lanes 1 and 2) or not (lane 3). To monitor the sensitivity of the hybridization, increasing amounts of cosmid containing known amounts of mini-exons were added to slots 5 – 11 of the gel used in Figure 3. Panel C shows that the hybridization of 0.25 mini-exon copy per haploid genome can be detected. The 225-kb chromosomes must therefore contain <0.25 copy. Analogous results were obtained in three other experiments.

All VSG mRNAs analysed thus far contain the mini-exon sequence at their 5′ end. Figure 3D shows that the mRNAs for VSG 1.1A and 1.3A are no exception. In this experiment total RNA was hybridized with AnTat 1.1A or 1.3A cDNA and the resulting hybrid was probed with a labeled mini-exon probe. Only homologous cDNA-RNA combinations gave positive results in this sandwich blot experiment.

These experiments show that the 225-kb chromosomes can yield VSG mRNAs with mini-exon sequence even though they contain no complete copy of this sequence. This result can be interpreted in two ways. Small chromosomes might contain split

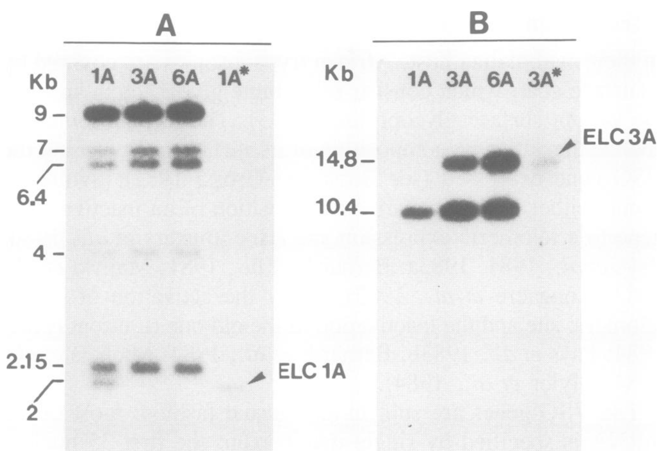


**Fig. 2.** Autoradiograms of the blot analysis of the 1.1A and 1.3A VSG gene present in total genomic DNA and in DNA from a 225-kb chromosome. The 225-kb chromosome of variant AnTat 1.1A (1A* in **panel A**) and variant AnTat 1.3A (3A* in **panel B**), were electroeluted and digested to completion with endonuclease *Pst*I and *Sph*I, respectively. The DNA was size-fractionated by agarose gel electrophoresis and run in parallel with 1 μg of the total DNA of variants 1.1A, 1.3A and 1.6A, each cut by either *Pst*I **(A)** or *Sph*I **(B)**. The DNA was blotted onto nitrocellulose filters and probed with a 750-bp *Eco*RI-*Pst*I fragment of an AnTat 1.10A cDNA in **panel A**. (The AnTat 1.10A probe is partly homologous to the AnTat 1.1A cDNA and therefore diminished hybridization with the 1.1A ELC fragment.) The 1.10A probe was used here because it hybridizes to all members of the 1.1 family (Pays *et al.*, 1983a). The blot in **panel B** was hybridized with a 1-kb *Pst*I insert of an AnTat 1.3A cDNA (Laurent *et al.*, 1983).

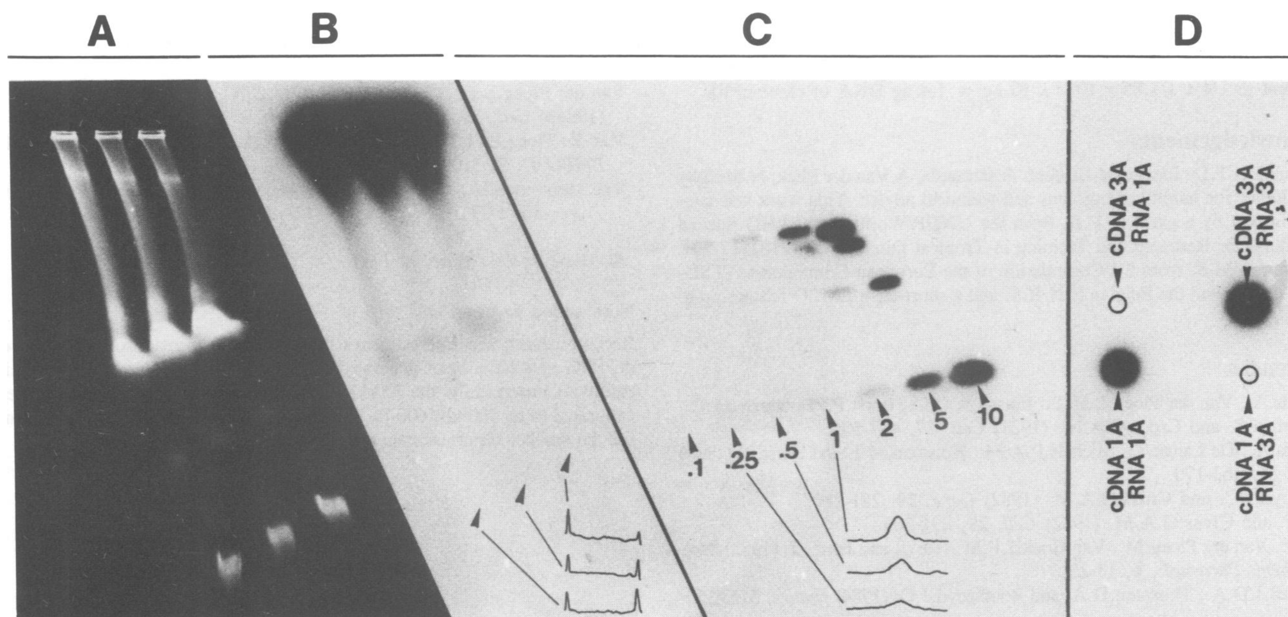


**Fig. 3.** Analysis of mini-exon sequences in 225-kb chromosomes and in the VSG mRNAs transcribed from these chromosomes. **Panels A, B** and **C** refer to the same PFG gel, the slots of which have been loaded with AnTat 1.1A, 1.3A and 1.6A variants (slots 1 – 3, respectively), and with increasing amounts of a *Pvu*II digest of the cosmid recombinant clone cPRI (slots 5 – 11). The latter cosmid contains four copies of the mini-exon sequence on a 0.7-kb fragment (arrowheaded), as well as additional copies on larger fragments (De Lange *et al.*, 1983). As detailed in Materials and methods, the amounts of cPRI DNA used correspond to a range of 0.1 – 10 mini-exon copies. **Panel A** shows the ethidium bromide staining of the first three lanes only, since the amounts of cPRI DNA used are too low to be visible. **Panels B** and **C** show the autoradiogram obtained following hybridization of a Southern blot of the PFG gel with a synthetic $^{32}$P-labeled 22-mer oligonucleotide complementary to part of the mini-exon sequence (De Lange *et al.*, 1983). The arrowheads in **panel B** indicate the position of the 225-kb chromosome. The regions of interest of the autoradiogram were scanned, along the electrophoretic migration axis, with a Joyce-Loebl micro-densitometer and the most relevant tracings are shown in **panels B** and **C**. The small peaks at both ends of the tracings in **panel B** correspond to pencil marks drawn symmetrically above and below the site of the 225-kb chromosome. **Panel D** shows an autoradiogram of a sandwich dot-blot hybridization, in which RNA selected by hybridization to the cloned cDNAs of AnTat 1.1A (Pays *et al.*, 1983a) and AnTat 1.3A (Laurent *et al.*, 1983) was hybridized to the labeled mini-exon 22-mer probe.

mini-exons, too short for detection by mini-exon probes, and which give rise to the mini-exon containing mRNAs from these chromosomes. The alternative is that trypanosomes can join RNAs made on different chromosomes. Whereas the first interpretation is far-fetched, the second is in full agreement with the results of the analysis of stable and nascent transcripts of mini-exon genes and the DNA repeats in which they are embedded (Kooter *et al.*, 1984; Campbell *et al.*, 1984; Milhausen *et al.*, 1984). The mini-exon has been found in the genome of several species of trypanosomes (De Lange *et al.*, 1983, 1984b; Milhausen *et al.*, 1984; Nelson *et al.*, 1984). Hence, we conclude that transcription of VSG genes and possibly all protein-coding genes in Kinetoplastida is discontinuous.

## Materials and methods

*Trypanosomes*

The trypanosome clones AnTat 1.1A, AnTat 1.3A and AnTat 1.6A were isolated from *T. brucei brucei*, strain EATRO (East African Trypanosomiasis Research Organisation) 1125, and have been described by Van Meirvenne *et al.* (1975). Trypanosomes were grown in rats and purified from blood elements on DEAE-cellulose column chromatography (Lanham and Godfrey, 1970).

*Pulsed field gradient (PFG) gel electrophoresis*

PFG gel electrophoresis (Schwartz *et al.*, 1983, 1984) of intact chromosomes was carried out as described (Van der Ploeg *et al.*, 1984a): $10^8$ lysed trypanosomes in agarose blocks were placed in a 1% agarose gel run at 20°C, alternating the electrical fields (17.5 V/cm North-South; 6.25 V/cm West-East) every 35 s for 18 h. The diagonal migration of the DNA is due to the combination of electrical fields applied alternately in the N/S and W/E direction. The mol. wts. determined with PFG gel electrophoresis are approximate. We have previously calibrated the size of the DNA molecules of *T. brucei*, strain 427 (Van der Ploeg *et al.*, 1984b) against yeast chromosomes and phage DNAs and we have used strain 427 DNA molecules as size markers in the experiments described here.

*Southern blotting and hybridization*

After PFG gel electrophoresis the DNA was transferred to a nitrocellulose filter (Southern, 1975). The PFG gels were pre-incubated with 0.25 N HCl for 30 min prior to the alkali denaturation step to reduce the mol. wt. of the DNA.

Filter hybridizations with nick-translated (Rigby *et al.*, 1977) cDNA restriction fragments isolated from low melting agarose gels were at 65°C in 3 x SSC (1 x SSC: 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) (Jeffreys and Flavell, 1977) with the addition of 10% dextran sulphate. Post-hybridizational washes were carried out at 65°C with several changes of 0.1 x SSC.

The 22-mer mini-exon probe was labeled at the 5′ end using polynucleotide kinase with [$^{32}$P]ATP (Maxam and Gilbert, 1980). Filter hybridizations with this probe were performed at 30°C in 6 x NET hybridization buffer (0.9 M NaCl; 6 mM EDTA.; 90 mM Tris-HCl, pH 7.5; 5 x Denhardt; 0.5% SDS) as described (De Lange *et al.*, 1983). Post-hybridizational washes were performed at the hybridization temperature in 3 x SSC, 0.1% SDS.

The sandwich hybridization described (Dunn and Hassel, 1977; Payne *et al.*, 1982) was modified (Dr T.De Lange, personal communication) as follows: 0.7 μg plasmid DNA was heated to 100°C in 3 μl of $H_2O$ and spotted onto nitrocellulose filters. These filters were hybridized to 1 mg total *T. brucei* RNA for 16 h at 65°C in 3 x SSC as described above. Post-hybridizational washes were carried out in 0.3 x SSC at 65°C for 2 h, following by pre-incubation (30 min) in the same hybridization mix without RNA and by hybridization to the labeled 22-mer.

*Purification of the chromosomes*

Chromosomes were isolated from preparative PFG agarose gels after ethidium bromide staining and purified from agarose by electro-elution in dialysis tubing (Van der Ploeg *et al.*, 1984b). The electro-elution was done in the PFG apparatus for 4 h under the conditions used for the preparative gels. The DNA was recovered from the buffer after repeated phenol extractions and ethanol precipitation.

*'Mini-exon' titration*

The reference is a 0.7-kb *Pvu*II restriction fragment from cosmid cPRI, which contains four copies of the mini-exon (De Lange *et al.*, 1983).

The *Pvu*II digest of cosmid cPRI was added on the gel 4 h before terminating electrophoresis. Since every *T. brucei* lane contains 10 μg DNA, we used 0.14 – 14 ng of cosmid clone DNA to quantitate the hybridization intensity of 0.1 – 10

mini-exon copies in the 0.7-kb *Pvu*II band. The size of the diploid (Tait, 1980) genome of *T. brucei* is 74 x $10^3$ kb (Borst *et al.*, 1982) which is 1.85 x $10^3$ the size of clone cPRI. One mini-exon copy in 10 $\mu g$ nuclear DNA is therefore equivalent to 1/4 x 1/1.85 x $10^{-3}$ x 10 $\mu g$ = 1.4 ng DNA of clone cPRI.

## Acknowledgements

We thank Drs T.De Lange, M.Laurent, A.Bernards, A.Van der Blick, N.Murphy and Y.Jooter for helpful discussions and technical advice. This work was supported in part by a grant to P.B. from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (T16/181/T7/34), by a grant to M.S. from the Commission of the European Communities (TSD-M-023-B(RS)) and the Belgian F.N.R.S. and a short-term EMBO fellowship to M.G.

## References

Bernards,A., Van der Ploeg,L.H.T., Frasch,A.C.C., Borst,P., Boothroyd,J.C., Coleman,S. and Cross,G.A.M. (1981) *Cell*, **27**, 497-505.

Bernards,A., De Lange,T., Michels,P.A.M., Huisman,M.J. and Borst,P. (1984) *Cell*, **36**, 163-170.

Boothroyd,J.C. and Cross,G.A.M. (1982) *Gene*, **20**, 281-289.

Borst,P. and Cross,G.A.M. (1982) *Cell*, **29**, 291-303.

Borst,P., Van der Ploeg,M., Van Hoek,J.F.M., Tas,J. and James,J. (1982) *Mol. Biochem. Parasitol.*, **6**, 13-23.

Campbell,J.D.A., Thornton,D.A. and Boothroyd,J.C. (1984) *Nature*, **311**, 350-355.

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., Veeneman,G.H., Cornelissen,A.W.C.A. and Borst,P. (1984a) *Nucleic Acids Res.*, **12**, 3777-3790.

De Lange,T., Berkvens,T.M., Veerman,H.J.G., Frasch,A.C.C., Barry,J.D. and Borst,P. (1984b) *Nucleic Acids Res.*, **12**, 4431-4443.

Dorfman,D.M. and Donelson,J.E. (1984) *Nucleic Acids Res.*, **12**, 4907-4920.

Dunn,A.R. and Hassel,J.A. (1977) *Cell*, **12**, 23-36.

Hoeijmakers,J.H.J., Frasch,A.C.C., Bernards,A., Borst,P. and Cross,G.A.M. (1980) *Nature*, **284**, 78-80.

Jeffreys,A.J. and Flavell,R.A. (1977) *Cell*, **12**, 429-439.

Kooter,J.M., De Lange,T. and Borst,P. (1984) *EMBO J.*, **3**, 2387-2392.

Lanham,S.M. and Godfrey,D.G. (1970) *Exp. Parasitol.*, **28**, 521-534.

Laurent,M., Pays,E., Magnus,E., Van Meirvenne,N., Matthyssens,G., Williams, R.D. and Steinert,M. (1983) *Nature*, **302**, 263-266.

Laurent,M., Pays,E., Delinte,K., Magnus,E., Van Meirvenne,N. and Steinert,M. (1984) *Nature*, **308**, 370-373.

Longacre,S., Hibner,U., Raibaud,A., Eisen,H., Baltz,T., Giroud,C. and Baltz,D. (1983) *Mol. Cell. Biol.*, **3**, 399-409.

Majiwa,P.A.O., Young,J.R., Englund,P.T., Shapiro,S.Z. and Williams,R.O. (1982) *Nature*, **297**, 514-516.

Maxam,A. and Gilbert,M. (1980) *Methods Enzymol.*, **65**, 499-560.

Michels,P.A.M., Van der Ploeg,L.H.T., Liu,A.Y.C. and Borst,P. (1984) *EMBO J.*, **3**, 1345-1351.

Milhausen,M., Nelson,R.G., Sather,S., Selkirk,M. and Agabian,N. (1984) *Cell*, **38**, 721-729.

Myler,P., Nelson,R.G., Agabian,N. and Stuart,K. (1984) *Nature*, **309**, 282-284.

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., Barr,P. and Agabian,N. (1984) *Nature*, **308**, 665-667.

Parsons,M., Nelson,R.G., Stuart,K. and Agabian,N. (1984a) *Proc. Natl. Acad. Sci. USA*, **81**, 684-688.

Parsons,M., Nelson,R.G., Stuart,K., Watkins,K.D. and Agabian,N. (1984b) *Cell*, **38**, 309-316.

Payne,G.S., Bishop,J.M. and Varmus,H.E. (1982) *Nature*, **295**, 209-213.

Pays,E., Van Meirvenne,N., Le Ray,D. and Steinert,M. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2673-2677.

Pays,E., Van Assel,S., Laurent,M., Darville,M., Vervoort,T., Van Meirvenne,N. and Steinert,M. (1983a) *Cell*, **34**, 371-381.

Pays,E., Delauw,M.F., Van Assel,S., Laurent,M., Vervoort,T., Van Meirvenne, N. and Steinert,M. (1983b) *Cell*, **35**, 721-731.

Rigby,P.W.J., Dieckmann,M., Rhodes,C. and Berg,P. (1977) *J. Mol. Biol.*, **113**, 237-251.

Schwartz,D.C. and Cantor,C.R. (1984) *Cell*, **37**, 67-75.

Schwartz,D.C., Saffran,W., Welsh,J., Haas,R., Goldenberg,M. and Cantor,C.R. (1983) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 189-195.

Southern,E. (1975) *J. Mol. Biol.*, **98**, 503-517.

Tait,A. (1980) *Nature*, **287**, 536-538.

Van der Ploeg,L.H.T., Liu,A.Y.C., Michels,P.A.M., De Lange,T., Borst,P., Majumder,H.K., Weber,H. and Van Boom,J.H. (1982) *Nucleic Acids Res.*, **10**, 3591-3604.

Van der Ploeg,L.H.T., Schwartz,D.C., Cantor,C.R. and Borst,P. (1984a) *Cell*, **37**, 77-84.

Van der Ploeg,L.H.T., Cornelissen,A.W.C.A., Michels,P.A.M. and Borst,P. (1984b) *Cell*, **39**, 213-221.

Van der Ploeg,L.H.T., Cornelissen,A.W.C.A., Barry,J.D. and Borst,P. (1984c) *EMBO J.*, **3**, 3109-3115.

Van Meirvenne,N., Janssens,P.G. and Magnus,E. (1975) *Ann. Soc. Belge Méd. Trop.*, **55**, 1-23.

*Received on 28 December 1984*

### Note added in proof

Recently, more accurate estimates for the sizes of trypanosome chromosomes on PFG gels have been generated using $\lambda$ ladders (A.Bernards, unpublished results). Consequently, the 700-kb, 1300-kb and 2-Mb bands have now been re-calibrated to be 500-kb, 600-kb and >700-kb, respectively. The size estimates for the smaller chromosomes remain unchanged.