MOLECULAR AND CELLULAR BIOLOGY, Jan. 1987, p. 357–364
0270-7306/87/010357-08$02.00/0
Copyright © 1987, American Society for Microbiology
Vol. 7, No. 1

Frequent Independent Duplicative Transpositions Activate a Single VSG Gene

MARY GWO-SHU LEE AND LEX H. T. VAN DER PLOEG*

Department of Genetics and Development, College of Physicians and Surgeons of Columbia University, New York, New York 10032

Received 26 June 1986/Accepted 7 October 1986

The expression of several surface antigen genes in *Trypanosoma brucei* is mediated by the duplicative transposition of a basic-copy variant surface glycoprotein (VSG) gene into an expression site. We determined that the appearance of variant 118, in a parasitemia, resulted from at least four independent duplicative transpositions of the same VSG 118 gene. Variants 117 and 118 both appeared at specific periods but resulted from multiple independent activations. Antigenic variants thus occur in an ordered manner. We show that in the duplicative transpositions of VSG genes, the ends of the transposed segments were homologous between the basic copy and the expression site. Sequences other than the previously reported 70-base-pair (bp) repeats could be involved. In one variant, 118 clone 1, the homology was between a sequence previously transposed into the expression site and a sequence located 6 kilobases upstream of the VSG 118 gene. In variant 118b the homology was presumably in 70-bp repeat arrays, while in a third 118 variant yet another sequence was involved. The possibility that the 70-bp repeats are important in the initial steps of the recombinational events was illustrated by a rearrangement involving a 70-bp repeat array. The data provide strong evidence for the notion that gene conversion mediates the duplicative transposition of VSG genes. We discuss a model that explains how the process of duplicative transposition can occur at random and still produce an ordered appearance of variants.

Trypanosoma brucei is a protozoan parasite transmitted by an insect vector, the tsetse fly, to a mammalian host. When living in the bloodstream of the mammalian host, *T. brucei* is covered by a dense protein coat consisting of a single glycoprotein, the variant surface glycoprotein (VSG; 12), which protects it against complement-mediated lysis. Moreover, by repeatedly changing the antigenic identity of the VSG coat, *T. brucei* escapes antibody-mediated immunodestruction (13, 38, 39). Each VSG is separately encoded, and there may be up to 1,000 different silent basic-copy VSG genes in the genome. The expression of these VSG genes occurs through different DNA recombinational processes, which translocate VSG genes to one of several telomerically located expression sites, creating an expression-linked copy (ELC; 14). It is this new ELC which is then transcribed (for a review, see P. Borst, Annu. Rev. Biochem., in press). Several other VSG genes that are always located at telomeres seem to be activated in situ (9, 19, 24, 34, 35, 41, 42).

A major route of VSG gene activation involves the duplicative transposition of a preexisting basic-copy (BC) gene into an expression site, presumably by unidirectional gene conversion (10, 26, 27, 33). All of the transposed segments analyzed thus far contain the VSG gene coding sequence and 1 to 2 kilobases (kb) of DNA located 5′ of it. The 5′ ends of the transposed segments are located within an imperfect 70-base-pair (bp) repeat (10, 18; T. De Lange, J. Luirink, J. M. Kooter, and P. Borst, EMBO J., in press). At the 3′ side, the transposed segment ends variably, either within the conserved C terminus or beyond the conserved 3′ end of the gene (6, 28).

Even though different mechanisms apply to the activation of VSG genes, some variants always arise early while others arise late. Antigenic variants may therefore be generated in order, as in the case of *T. equiperdum* (11, 23, 40; A. Y. C. Liu, P. A. M. Michels, A. Bernards, and P. Borst, J. Mol. Biol. in press). We wanted to determine the mechanism of duplicative transposition and identify whether variants arise in order.

* Corresponding author.

MATERIALS AND METHODS

Strains and cloning of trypanosomes. *T. brucei* variant 121 has been described by Liu et al. (J. Mol. Biol., in press), and variant 118b by Michels et al. (21, 22). Trypanosomes were cloned as described by Michels et al. (22).

DNA isolation and Southern blotting analysis. *T. brucei* was grown in rats and purified from blood elements as described by Fairlamb et al. (16). Nuclear DNA was isolated from trypanosomes as described by Van der Ploeg et al. (33) and digested with a fourfold excess of restriction endonuclease, and the products were resolved on 0.8% agarose gels. Following transfer to nitrocellulose (32), the filters were prehybridized and hybridized with ^{32}p-labeled probes (17, 29). After hybridization, filters were washed to a final stringency of 0.2 × SSC–0.1% sodium dodecyl sulfate (SDS) at 68°C.

Construction and screening of genomic libraries. Nuclear DNAs of variant 118 clones 1 and 4 were completely digested with *Eco*RI and *Hin*dIII, respectively, and then ligated into the plasmid vector pUC18. These ligation mixtures were used to transform *Escherichia coli* HB101. About 2×10^4 transformants were screened for each library with a ^{32}P-labeled probe of VSG 118 (5′ *Eco*RI-*Pst*I fragment from TcV 118-2) and the 1.8-kb *Hin*dIII fragment described in Fig. 5. The isolated genomic clones were then used for subsequent characterizations.

RNA preparation and analysis. White male Wistar rats (300 g) were used to grow trypanosomes for RNA and DNA isolations. Total RNA was isolated by LiCl precipitation described by Auffray and Rougeon (2). Northern analysis of

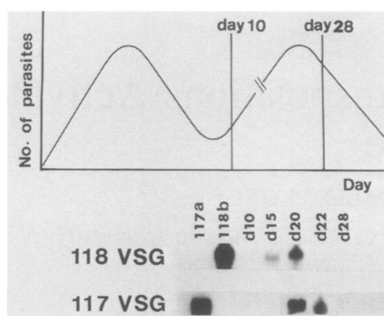

FIG. 1. Ordered appearance of antigenic variants in *T. brucei*. A cloned population of variant 121 was used to initiate a chronic infection of a rabbit. Blood samples were collected at daily intervals; the parasites were amplified in rats, and RNA was isolated, transferred to nitrocellulose, and hybridized with VSG 118 and 117 5'-specific cDNAs. (Top) Schematic representation of the parasitemia; (bottom) dot blot Northern analysis of the samples. Lanes 117a and 118b contain control RNAs from variants 117a and 118b, respectively. Lanes d10 through d28 contain RNA isolated from the mixed populations on the indicated days.

RNA was carried out on 1% agarose gels containing 2.2 M formaldehyde as described previously (8). The blots were hybridized overnight in 50% formamide–5× Denhardt solution–3× SSC–20 mM phosphate, pH 7.0, at 42°C with ^{32}P-labeled probes. Following hybridization, filters were washed to a final stringency of 0.2× SSC–0.1% SDS at 68°C.

DNA sequence analysis. Sequence analysis was performed by the dideoxy chain termination method (30). Two approaches were used to generate suitable DNA fragments for subcloning into the replicative forms of bacteriophage M13, mp18 and mp19. In the first method, DNAs were cleaved with restriction enzymes and ligated into the appropriate M13 vector. In the second method, DNA subclones were treated with Bal 31 exonuclease for various times at room temperature to generate a series of overlapping clones. These were subsequently ligated into the M13 vectors.

Pulsed-field gel electrophoresis. The electrophoresis units and DNA samples were as described by Schwartz and Cantor (31) and Van der Ploeg et al. (35). Agarose blocks containing trypanosome chromosome-sized DNA molecules were placed in the slots of a 1.0% agarose gel. DNA was electrophoresed for 22 h at a pulse frequency of 20 s at 210 V/cm at 20°C. After electrophoresis the DNA was stained with ethidium bromide and transferred to nitrocellulose as described previously. The filter was hybridized with a ^{32}P-labeled cDNA probe (500-bp *Eco*RI-*Pst*I fragment of TcV 118-2 [6]). Following hybridization, the filter was washed to a final stringency of 0.1× SSC at 68°C.

RESULTS

The appearance of a single variant occurs through the generation of different ELCs. We determined whether the timed appearance of a single new variant results from many independent activations of the same VSG gene. This would provide strong evidence for the notion that an order exists in which antigenic variants arise. A rabbit was infected with a clonal population of *T. brucei* variant 121. We then searched for the appearance of variants 118 and 117 to determine the number of independent activation events which led to the presence of these variants. At daily intervals through day 50, blood samples were collected. All samples were tested with fluorescent antibodies for the appearance of variant 118. In the infected rabbit, variant 118 was detected at days 15 through 22, with a maximum of 30% positive fluorescing variants at day 20 (Fig. 1). RNA samples were obtained from parasite populations at days 10, 15, 20, 22, and 28 of the parasitemia, transferred onto nitrocellulose filters, and probed with cDNAs of VSG 117 and VSG 118 (Fig. 1). Transcription of the 118 VSG gene was detected from days 15 through 22. Transcription of the 117 VSG gene was detected after day 20.

Because the BC of VSG gene 118 is present as a single copy, located at an internal position in the genome, the activation of the 118 VSG gene must occur by duplicative transposition of the gene into an expression site (37). Therefore, the number of different ELCs which are detected in the population may represent the minimal number of individual activations. We analyzed the DNA from the heterogeneous population of parasites isolated on day 20 for the presence of VSG 118 ELC fragments. The VSG 118 cDNA probe hybridized to the 5' end of the gene only in *Eco*RI-digested nuclear DNA. It therefore does not detect length alterations of the ELC which result from growth or deletions of the telomeric repeats. A total of at least six distinct 118 ELCs were detected in the *Eco*RI-digested DNA of the day 20 mixed population (Fig. 2, lane d20; as determined from a comparison of lane d20 with the other lanes). Several of these ELCs were already present in the day 15 population, in which we first identified variant 118 (data not shown). That

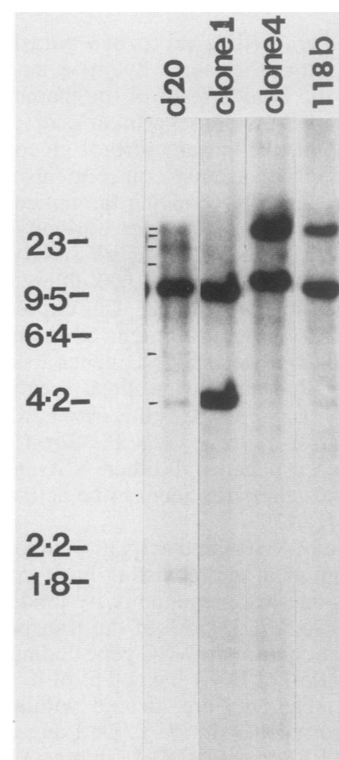

FIG. 2. Analysis of VSG 118 ELC fragments. Nuclear DNA from a heterogeneous population of trypanosomes obtained from day 20 (d20) of a parasitaemia, variant 118 clone 1 and clone 4, and variant 118b were digested with *Eco*RI, size-fractionated on 0.8% agarose gels, and transferred to nitrocellulose filters. The filter was hybridized with a ^{32}P-labeled cDNA probe (*Pst*I-*Eco*RI fragment from the 5' end of cDNA TcV 118-2 [6]. The different 118 ELC fragments in the sample isolated from the day 20 parasitemia are indicated with lines. Sizes are indicated in kilobases).

these extra fragments in the mixed population were indeed ELCs can be seen from the hybridization of three cloned VSG 118 variants obtained from the day 20 mixed population (Fig. 2). These ELCs were located on at least two different chromosomes, of 2,000 and 340 kb, as determined by pulsed-field gradient gel electrophoresis (Fig. 3, arrowheads).

Independent duplicative transpositions of VSG 118. To better understand the mechanism of duplicative transposition and to prove that the different ELCs resulted from independent VSG gene activations, we compared different cloned variants obtained from the day 20 population of the parasitemia. Three VSG 118-expressing variants were compared: 118 clone 1, 118 clone 4, and variant 118b. Each clone in Fig. 2 had a 118 ELC which corresponded in mobility to a band in the Southern blot containing the DNA from the day 20 population; furthermore, all three variants transcribed the VSG 118 gene, as determined by Northern analysis (see Fig. 8). To determine the size of the transposed segment and the region of the breakpoint in the 118 ELC of clones 1 and 4, nuclear DNAs isolated from these clones were digested with restriction endonucleases, transferred onto nitrocellulose, and hybridized with three different probes isolated from the 118 BC gene (Fig. 4, see legend for mapping strategy). Based on these hybridization patterns, physical maps of the two clones were constructed. Figure 5 shows the physical maps of the ELC genes of all three independently cloned VSG 118 variants; variant 118b has been published previously (20) and has a transposed segment of about 4 kb, while variant 118 clone 1 and variant 118 clone 4 both had transposed segments of about 8 kb. Moreover, the clone 4 transposed segment was different from that of clone 1 in that it had 70-bp repeat insertions in the 1.0-kb *Pvu*II fragment of both the BC and ELC (see below).

The presence of an insertion element in both the BC and ELC of clone 4 shows that its ELC arose independently of clone 1 and variant 118b. Also, the differences in the length of the transposed segments and the differences in the physical maps between clones 1 and 4 and variant 118b demonstrate that all three ELCs resulted from independent duplicative transpositions. Since these three ELCs were located in the 2,000-kb chromosomes, as shown by pulsed-field gel electrophoresis (data not shown), they probably represent one expression site only. However, in the day 20 parasite population we found hybridization of the VSG 118 cDNA

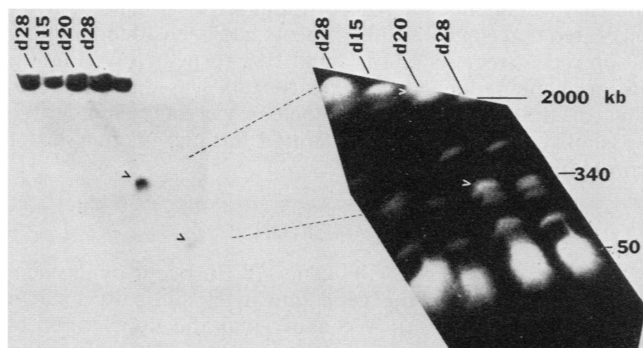

FIG. 3. Chromosomal location of the VSG 118 ELCs. Agarose blocks containing trypanosome DNA samples isolated at day 15, day 20, and day 28 of the parasitemia were loaded in the s.ots of a 1.0% agarose gel and electrophoresed for 22 h with a pulse frequency of 20s. The gel was stained with ethidium bromide (right-hand panel), and the DNA was transferred to nitrocellulose and then hybridized with the ^{32}P-labeled 500-bp *Pst*I-*Eco*RI cDNA fragment from the 5' end of clone TcV 118-2 (left-hand panel).

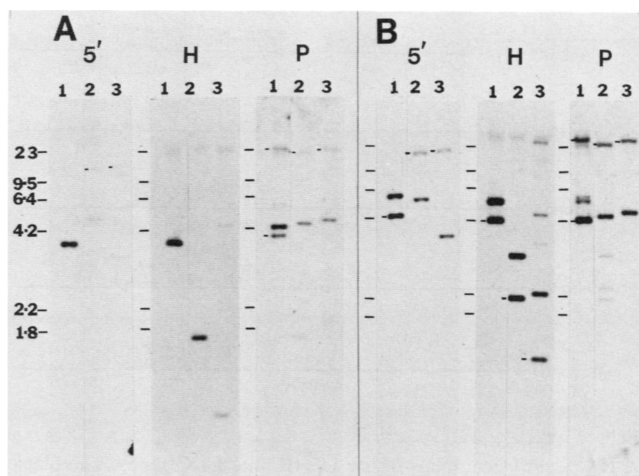

FIG. 4. Characterization of the 118 ELC and BC genes in variant 118 clone 1 and clone 4. The physical maps and location of probes are indicated in Fig. 5. Digestion was with *Bgl*II (lanes 1), *Hin*dIII (lanes 2), *Pvu*II (lanes 3). (A) The transposed segment of clone was about 8 kb in length, which is much larger than previously analyzed transposed segments. The length of transposed segment is illustrated by the hybridizations of *Bgl*II-digested DNA (lanes 1) with probes that are derived from an area up to 6.0 kp upstream of the BC gene, which all detect ELC fragments. Because *Bgl*II cuts within the 118 gene, the 5' cDNA probe (panel 5') and probe H (panel H) (these probes span an area up to 2.8 kb upstream of the gene) detect a single 3.9-kb fragment containing both the ELC and BC fragments. Only in the hybridization with a probe located even further upstream (probe P) did we detect the end of the transposed segment. In this case the 3.9-kb *Bgl*II ELC and BC fragments, a 4.2-kb BC *Bgl*II fragment, and a 23-kb *Bgl*II ELC fragment hybridized (probe P, lane 1). The faint band at 6 kb in lane 1 was common to all variants and therefore unrelated to the 118 VSG gene transposition. The telomere end of the ELC was mapped with the 5' cDNA probe in *Hin*dIII and *Pvu*II digests (lanes 2 and 3, respectively). (B) The transposed segment in clone 4 was also large (about 8 kb). However, compared with clone 1, we found different BC and ELC fragments in clone 4. The 5' cDNA probe (panel 5') and probe H (panel H) both detected the same two *Bgl*II fragments of 4.2 and 5.4 kb (lanes 1). The BC copy fragment in clone 4 was 1.5 kb larger than that of clone 1, and the ELC fragment was 0.3 kb larger. This was confirmed by analysis of BC and ELC clones that contained these rearranged sequences. The presence of inserted stretches of DNA in the 70-bp repeat array of the BC and ELC of clone 4 is illustrated in the hybridization with probe H. The clone 1 1.8-kb *Hin*dIII fragment, containing both the BC and ELC are respectively 3.3 kb and 2.1 kb in length in clone 4 (panel H, lane 2). In the *Hin*dIII and *Pvu*II digests we mapped ELC fragments containing the tip of the telomere (probe 5' lanes 2 and 3). With the 4.2-kb *Pvu*II probe (probe P) we determined the 5' end of the transposed segment; in the *Bgl*II digest (probe P, lane 1) we detected the ELC fragment of 4.2 kb, which comigrated with the 4.2-kb *Bgl*II BC fragment; in addition we detected the 5.4-kb enlarged BC fragment and the 6.0-kb fragment which was common to clone 1 and clone 4. The 25-kb *Bgl*II ELC fragment spanned the 5' end of the transposed segment. Hybridization with probe P also allowed mapping of the *Pvu*II and *Hin*dIII fragments located upstream in the expression site (ELC fragments indicated with dark lines). These physical maps were verified with the cloned and purified ELC fragments, 6.0-kb *Eco*RI fragment of clone 1, and 2.1-kb *Hin*dIII ELC fragment of clone 4.

with the 340-kb chromosome (Fig. 3). This indicates that duplicative transpositions occurred at other expression sites on different chromosomes. The appearance of variant 118 in this parasitemia therefore resulted from at least four (three at a 2,000-kb and one at a 340-kb chromosome) independent duplicative transpositions of the same VSG 118 gene. As at

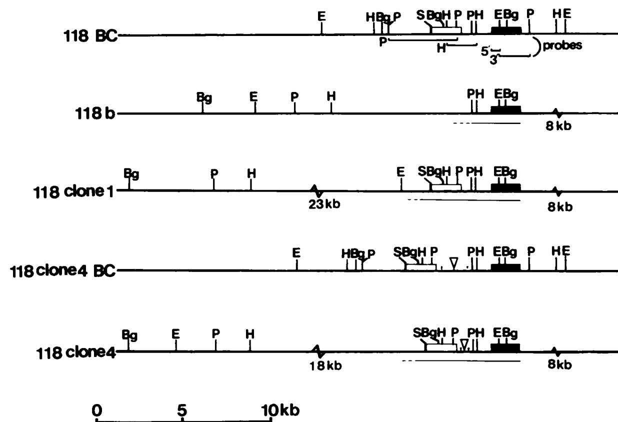

FIG. 5. Physical maps of the 118 BC and ELC genes. The black boxes indicate the coding sequences of VSG 118 gene. The open boxes indicate the coding region of VSG pseudogene MiTat 1.1000. The DNA fragments used as hybridization probes are shown below the map of the 118 BC. Probes used: 5', 500-bp *Pst*I-*Eco*RI cDNA fragment from the 5' end of clone TcV 118-2 (6); H, 1.8-kb *Hin*dIII fragment located within the transposed segment (it contains many 70-bp repeats, which cause a faint smear in the hybridizations with all variants in the high-molecular-weight range); P, 4.2-kb *Pvu*II fragment located upstream of the 70-bp repeats. The size of the transposed segment in each variant is indicated below the ELC maps with a line. The open triangle represents insertions, and the flanking arrows show the integration sites. Abbreviations: Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; P, *Pvu*II; S, *Sal*I.

least six distinct 118 ELCs could be detected in the day 20 population (Fig. 2), numerous simultaneous activations of the same VSG 118 gene must have occurred.

Duplicative transposition of VSG genes is mediated by homologous flanking sequences. Thus far, the 5' ends of the transposed segments have been located in an imperfect 70-bp repeat (10; De Lange et al., EMBO in press). Another case described previously (27) demonstrated the presence of a variant with an aberrant 5' end of the transposed segment. However, the 5' border sequence of this transposed segment remained unidentified. Because the expression sites contained 70-bp repeats, it can be assumed that these repeats function in aligning the BC and ELC site. Evidence for this notion can be obtained by establishing that the 5' borders of other transposed segments retain homology between BC and ELC but can consist of different DNA sequences. We therefore determined the 5' border sequence of the BC and ELC of a transposed segment which did not end in a 70-bp repeat. Unlike variant 118b (21, 22), the 5' border of the transposed segment in clone 1 and clone 4 was not located in the 70-bp repeat region.

To determine the nucleotide sequence of the 5' end of the transposed segment, we isolated a 6-kb *Eco*RI fragment from the 118 ELC of clone 1. The sequences around the break junction of both the BC and the ELC of the 118 VSG gene in clone 1 were determined (Fig. 6). The 5' end of the transposed segment in clone 1 was located in a stretch of DNA about 350 bp long, in which there was 60% homology between the BC and the ELC and which completely differed from the 70-bp repeats. In Fig. 6 this region is divided into two parts, a 50% homologous region and a region which retains 70% homology between the BC and the ELC (HR region). The sequence downstream of this homologous region was identical to that of the BC. The fragment which was about 50% homologous between the BC and the ELC (ES in Fig. 6) hybridized to two bands in a *Pvu*II digest of nuclear DNA from different variants. However, in variant 118 clone 1, a third band was visible (Fig. 7, arrowhead). This third hybridizing *Pvu*II fragment correponded to the VSG 118 ELC fragment of clone 1. Moreover, the ES sequence did not hybridize to the 4.2-kb *Pvu*II BC fragment of VSG 118, which contained the 5' end of the transposed segment. Because the ES sequence preexists in the genome, as deduced from its hybridization to two BC bands in nuclear DNA, it is unlikely that the ES sequence was created at random during duplicative transposition of VSG 118 by a sloppy DNA polymerase. The ES sequence must therefore have been duplicated and transposed into the expression site previous to the VSG 118 transposition. Because the ELC of variant 118 clone 4 did not hybridize with the ES probe, its 5' breakpoint must be located in another DNA sequence.

Therefore, different sequences (70-bp repeats and the 350-bp region presented here) that are homologous between the BC and ELC can flank the ends of the transposed segments. Their function may be to align the BC and the ELC. Apparently, homology between the transposed segment and its target site, as opposed to the presence of a particular nucleotide sequence, is essential for duplicative transposition.

Duplicative transposition of two linked VSG genes. The transposed segments of clones 1 and 4 both carried two linked VSG genes, the MiTat 1.1000 VSG gene, located proximally, and the 118 VSG gene, located distally. However, the mRNA of the 118 VSG gene was the only detectable mature VSG mRNA (Fig. 8). The presence of several stop codons in the nucleotide sequence of the MiTat 1.1000 gene showed that it was a pseudogene (Fig. 6, boxed sequences). The VSG 221 expression site also possesses two VSG genes at an expression site (5). However, both the VSG 221 and its pseudogene produced mature mRNAs. We are currently analyzing whether the MiTat 1.1000 pseudogene is transcribed in nuclear run-on assays.

70-bp repeats inserted at the BC and ELC of variant 118 clone 4. We found an insertion element in the 70-bp repeats located upstream of the 118 gene of clone 4 (Fig. 9). To identify conserved sequences that may mediate recombination at the 70-bp repeats, we analyzed these inserted sequences. The integration occurred between the fourth and fifth 70-bp repeats of the original BC; however, the length of the inserted fragment was 1,500 bp in the BC and 300 bp in the ELC. Sequence analysis of the inserted fragments demonstrated that 4 and 13 70-bp repeats had been added into the 70-bp repeat regions of ELC and BC, respectively (Fig. 9). Because the sequence of these repeats was different from that of the flanking 70-bp repeats, we can exclude the possibility that they were generated through unequal sister chromatid exchange.

DISCUSSION

All activations of the VSG gene 118 BC occur by duplicative transposition of the gene into a telomerically located expression site (4, 33). We show that the appearance of variant 118 occurs in an ordered manner, because many independent activations of the same VSG 118 gene were detected at one specific period of a parasitemia. Variant 118 appeared earlier than variant 117, and the occurrence of both variants was accompanied by the presence of many different ELCs (data for VSG 117 not shown). We have demonstrated at least three distinct duplicative transposition events of the VSG 118 gene, occurring at one expression site, located in

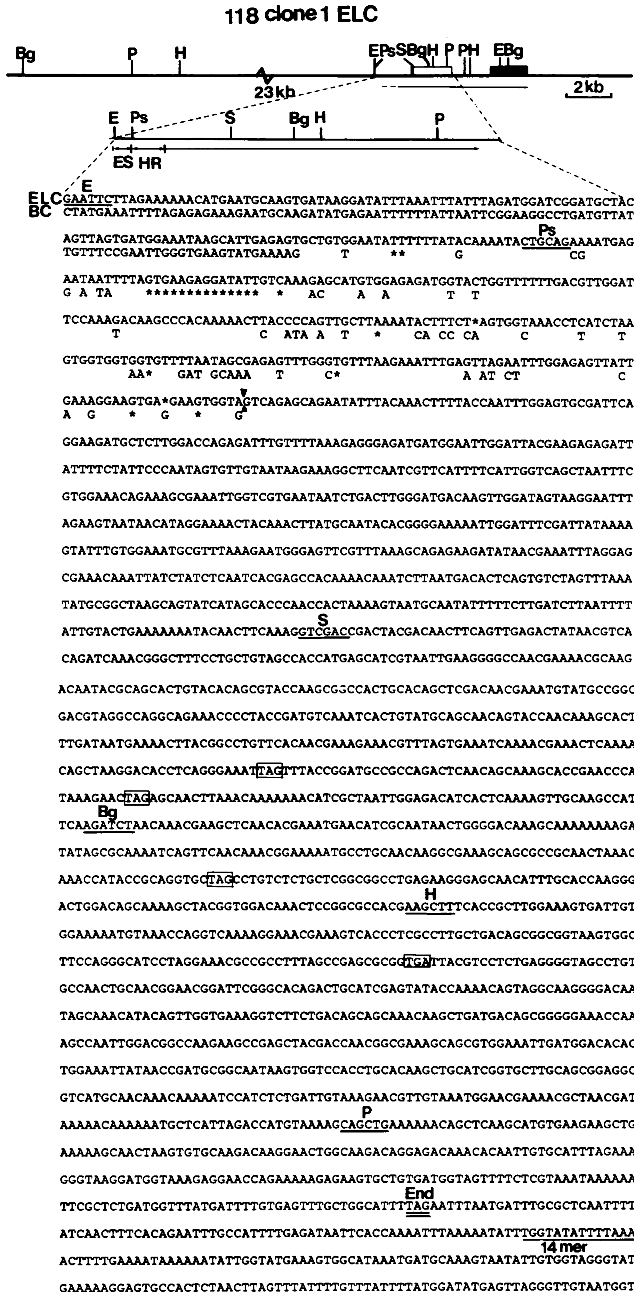

FIG. 6. Sequence of the 5′ border of the transposed segment and of the VSG pseudogene miTat 1.1000 in variant 118 clone 1. Nucleotide sequence was determined by the dideoxy chain termination method (30). The top line shows the sequence of the ELC; the lower line shows the differences in the nucleotide sequence of the BC; the first 100 nucleotides are 50% homologous between BC and ELC; the nucleotide sequence beyond this point only needs deletions in the BC strand for optimal alignment. The nucleotide sequence, except for the first 100 nucleotides, is aligned to optimize homology. ES, expression site-specific sequence between *Eco*RI and *Pst*I sites. This sequence is 50% homologous between the BC and the ELC. This sequence was used as a probe to detect its copy number in the trypanosome genome. HR represents the 70% homologous region between the BC and the ELC. Sequences downstream of HR are completely identical between the BC and the ELC. The arrows in the sequence indicate the starting position of the identical region. Stars show missing residues. The open boxes indicate the in-frame termination codons in the coding region of the MiTat 1.1000 gene. The double line shows the end of the coding region of the MiTat 1.1000 gene. Abbreviations: Bg, *Bgl*II; H, *Hin*dIII; S, *Sal*I. 14-mer, Conserved nucleotide sequence at the end of most VSG genes.

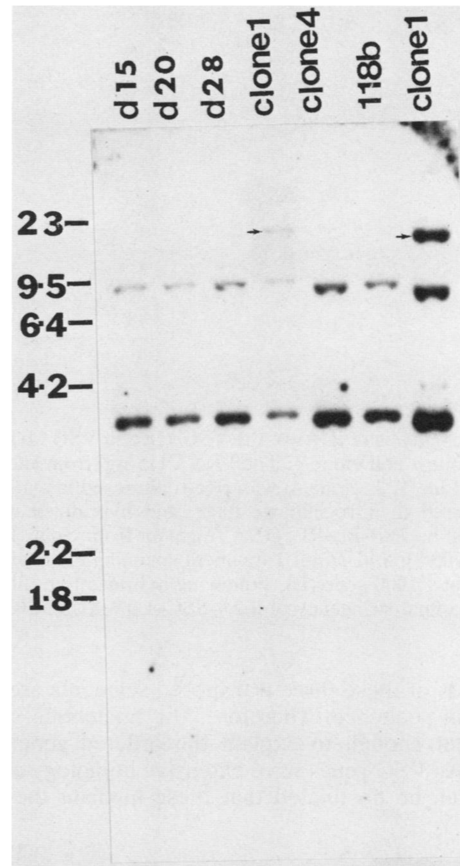

FIG. 7. Characterization of the ES sequence in variant 118 clone 1. Nuclear DNAs were digested with the restriction enzyme *Pvu*II, size-fractionated in 0.8% agarose gels, transferred to nitrocellulose, and hybridized with the 100-bp *Pst*I-*Eco*RI fragment of clone 1 shown in Fig. 6 as ES. Following hybridization, filters were washed to a final stringency of 0.2× SSC–0.1% SDS at 68°C. The origin of the DNA samples is indicated at the top.

the 2,000-kb chromosomes. However, transpositions into other expression sites occurred, as shown by the hybridization of VSG 118 cDNA with a 340-kb chromosome (Fig. 3). We have not shown that this gene copy on the 340-kb chromosome is transcribed. However, this ELC was present at day 20 and not in the day 28 population. Therefore, it must have been selected against. Since this selection probably resulted from the antibody-mediated killing of variant 118 trypanosomes, the VSG 118 gene may have been transcribed.

The data show that a timing mechanism must exist which allows the ordered appearance of variants. It is unlikely that the timing of expression is regulated by sequence homologies between the BC and the ELC. In variant 118 clone 1 the 5′ end of the transposed segment was located in a 350-bp stretch of DNA which was 60% homologous between the BC and the ELC. The 5′ end of the clone 4 transposed segment was located in yet another sequence, because its ELC did not hybridize with the clone 1 breakpoint sequence. The breakpoint of the previously described variant 118b is, as in most other variants, located in the 70-bp repeat array (21).

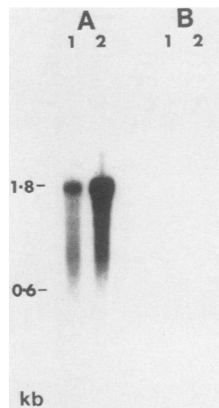

FIG. 8. Analysis of RNA of the VSG 118 and VSG MiTat 1.1000 genes in clone 1 and clone 4. Total RNA (15 μg) from each variant (lanes: 1, clone 1; 2, clone 4) was electrophoresed in a 1% agarose gel transferred to nitrocellulose filter, and hybridized with a 5′-oriented 500-bp *Pst*I-*Eco*RI cDNA fragment from clone TcV 118-2 (A) or the 700-bp *Sal*I-*Hin*dIII fragment spanning the coding region of the MiTat 1.1000 gene (B). Following hybridization, filters were washed to a final stringency of 0.2× SSC–1.0% SDS at 68°C.

The 5′ ends of these three transposed segments are thus all different in sequence. Therefore, the nucleotide sequence itself is not enough to explain the ordered generation of variants. As VSG genes have extensive homology at their 3′ ends, it can be postulated that these mediate the ordered appearance. However, all VSG genes have related sequences at their C-terminal coding regions and in their 3′ untranslated regions; therefore, it is unlikely that these could determine order.

Another possibility is that frequent duplicative transpositions occur randomly, allowing any homologous sequence to mediate duplicative transposition. Specific switches, such as those we detected, would then have been selected for by another mechanism. Since the switch frequency estimated in vitro is too high (10^{-6} to 10^{-7} per division; G. S. Lamont, R. S. Tucker, and G. A. M. Cross, Parasitology, in press) to explain the limited number of different variants that occur in vivo, selection seems to be a real option. Selection could result from diminished viability of a short-lived switch intermediate (36). Variants with coats that consist of more than one type of VSG have been identified (3, 15). These may represent switch intermediates. Thus, if only a few combinations of VGSs in double coats are protective in vivo, successful antigenic switches will be few. The timed and independent activations of VSG gene 118 by different duplicative transpositions can thus be explained if they occurred in a common precursor variant.

Pays et al. (26) have provided evidence that segmental gene conversions among telomeric VSG genes can alter the VSG gene coding sequence. Moreover, based on the structure of the transposed segment, a process equivalent to the gene conversion which occurs during mating type switches in yeast has been proposed to mediate the duplicative transposition of VSG genes from a chromosomal internal location to a telomeric position (10, 25, 33). Duplicative

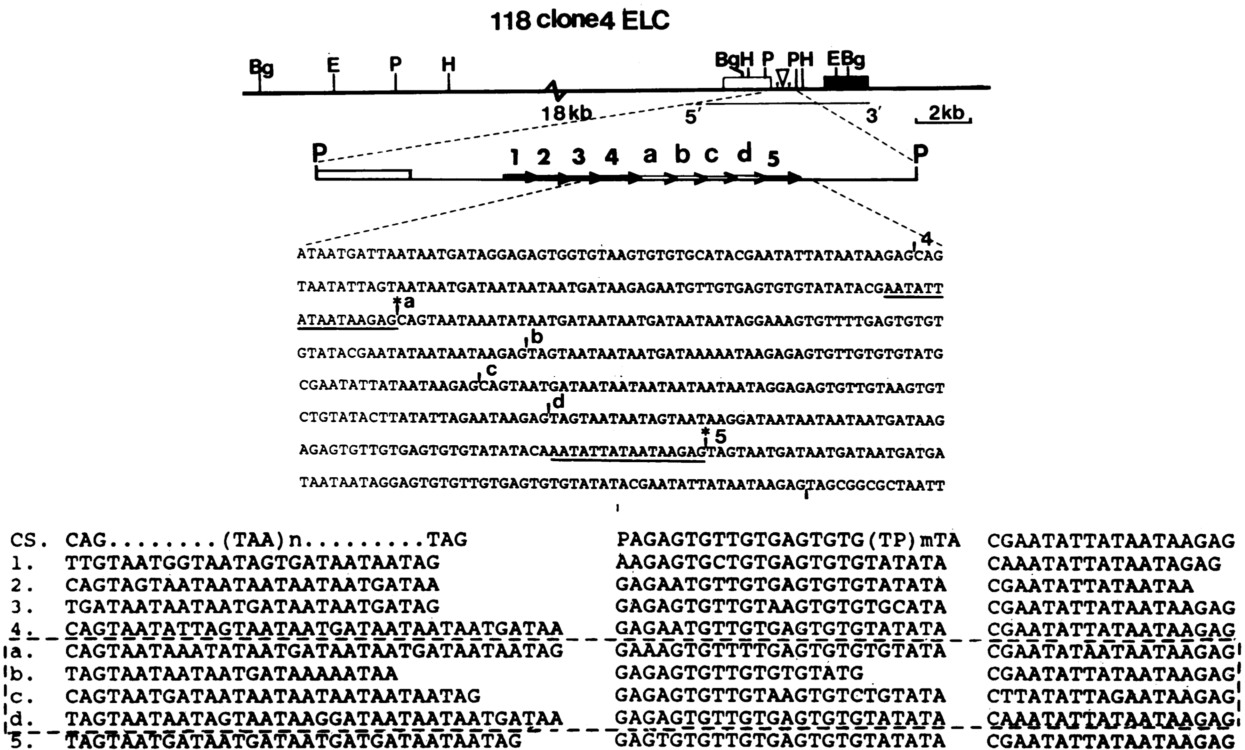

FIG. 9. Nucleotide sequence of the insertion in the ELC of variant 118 clone 4. Regions containing DNA of interest were subcloned and sequenced by the dideoxy chain termination method (30). (Top) Physical map of the VSG 118 clone 4 ELC. 1, 2, 3, 4, and 5 represent the 70-bp repeats in the original VSG 118 BC. a, b, c, and d represent the newly inserted 70-bp repeats in the 118 ELC of clone 4. The positions of the starting site of each repeat are marked by arrows. The location of the putative integration site is marked by asterisks. (Bottom) Comparison of the original 70-bp repeats (numbered 1 to 5) and the newly inserted 70-bp repeats (lettered a to d). CS, 70-bp repeat consensus sequence; n and m indicate variable numbers of TAA and TP nucleotides, respectively. P. Purine nucleotide. See the legend to Fig. 5 for other abbreviations.

transposition via gene conversion requires homologous pairing of the ends of the transposed segments with the target site. We have demonstrated that a large (at least 350 bp) stretch of DNA exists at the 5′ end of the transposed segment which is homologous between the BC and the ELC. This observation provides strong evidence for gene conversion as the mechanism for duplicative transposition of VSG genes. It also indicates that the 70-bp repeats that flank most transposed segments can be involved in aligning donor and target, but that the nucleotide sequence of the homologous region is not important.

One of the primary functions of the 70-bp repeat may be initiation of recombination. Insertion elements integrated at identical nucleotide positions in the BC and ELC of variant 118 clone 4 are indicative of 70-bp repeat specific recombinational events. The sequence analysis of the insert and flanking sequences excludes the possibility that unequal sister chromatid exchange created the insertions. A comparison of 70-bp repeat recombinations does not allow an accurate determination of their breakpoints, again due to the repetitive nature of the sequences (1, 4, 7, 10; De Lange et al., in press).

The inserts at the BC and ELC of variant 118 clone 4 were of different lengths. This length difference cannot be explained by independent transpositions into the same target site, because such integration events are rare and would not explain how integration of the same repeats occurred in front of the same gene at the same site. One possibility is that there was an initial insertion into the BC, which was rearranged after translocation to the ELC. Alternatively, it is possible that a single event occurred at duplicative transposition. This could be explained by integration of 70-bp repeats during the process of gene conversion. After separation of the BC and the ELC, two heteroduplex molecules then result which repaired differently, resulting in variation of the insert sizes. We are currently testing whether specific DNA-binding proteins are present that can recognize and perhaps cleave the 70-bp repeat and in this way initiate duplicative transposition.

ACKNOWLEDGMENTS

We thank David Misek, David Glass, Charles Cantor, Cathy Shea, Hans Ronne, and Ramona Polvere for critical reading of the manuscript. We thank P. Borst, Mark Timmers, and Titia De Lange for sharing unpublished results and for a gift of a heterogeneous trypanosome population from an infected rabbit.

This work was supported by Public Health Service grant AI-21784 from the National Institutes of Health and a Searle scholarship to L.H.T.V.D.P. The authors are supported by a grant from the John D. and Catherine T. MarcArthur Foundation.

LITERATURE CITED

1. **Aline, R., Jr., G. MacDonald, E. Brown, J. Allison, P. Myer, V. Rothwell, and K. Stuart.** 1985. (TAA)n within sequences flanking several intrachromosomal variant surface glycoprotein genes in *Trypanosoma brucei*. Nucleic Acids Res. **13:** 3161–3177.
2. **Auffray, C., and F. Rougeon.** 1980. Purification of mouse immunoglobulin heavy-chain RNAs from total myeloma tumor RNA. Eur. J. Biochem. **107:**303–314.
3. **Baltz, T., C. Giroud, D. Baltz, C. Roth, A. Raibaud, and H. Eisen.** 1986. Stable expression of two variable surface glycoproteins by cloned *Trypanosoma equiperdum*. Nature (London) **319:**602–604.
4. **Bernards, A., T. De Lange, P. A. M. Michels, A. Y. C. Liu, M. J. Huisman, and P. Borst.** 1984. Two modes of activation of a single surface antigen gene of *Trypanosoma brucei*. Cell **36:**163–170.
5. **Bernards, A., J. M. Kooter, and P. Borst.** 1985. Structure and transcription of a telomeric surface antigen gene of *Trypanosoma brucei*. Mol. Cell. Biol. **5:**545–553.
6. **Bernards, A., L. H. T. Van der Ploeg, A. C. C. Frasch, P. Borst, J. C. Boothroyd, S. Coleman, and G. A. M. Cross.** 1981. Activation of trypanosome surface glycoprotein genes involves a gene duplication-transposition leading to an altered 3′ end. Cell **27:**497–505.
7. **Bernards, A., L. H. T. Van der Ploeg, W. C. Gibson, P. Leegwater, F. Eijgenraam, T. De Lange, P. Weijers, J. Calafat, and P. Borst.** 1986. Rapid change of the repertoire for variant surface glycoprotein genes in trypanosomes by gene duplication and deletion. J. Mol. Biol. **190:**1–10.
8. **Boedtker, H.** 1971. Conformation-independent molecular weight determination of RNA by gel electrophoresis. Biochim. Biophys. Acta **240:**448–453.
9. **Borst, P., A. Bernards, L. H. T. Van der Ploeg, P. A. M. Michels, A. Y. C. Liu, T. De Lange, and J. M. Kooter.** 1983. The control of variant surfac antigen synthesis in trypanosomes. Eur. J. Biochem. **137:**383–398.
10. **Campbell, D. A., M. P. Van Bree, and J. C. Boothroyd.** 1984. The 5′-limit of transposition and upstream barren region of a trypanosome VSG gene: tandem 76 base-pair repeats flanking (TAA) 90. Nucleic Acids Res. **12:**2759–2774.
11. **Capbern, A., C. Giroud, T. Baltz, and P. Mattern.** 1977. Trypanosoma equiperdum: étude de variations antigénique au cours de la trypanosomose experimentale du lapin. Exp. Parasitol. **42:**6–13.
12. **Cross, G. A. M.** 1975. Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. Parasitology **71:**393–417.
13. **Cross, G. A. M.** 1978. Antigenic variation in trypanosomes. Proc. R. Soc. Lond. Ser. B **202:**55–72.
14. **De Lange, T., and P. Borst.** 1982. Genomic environment of the expression-linked extra copies of genes for surface antigens of *Trypanosoma brucei* resembles the end of a chromosome. Nature (London) **229:**415–453.
15. **Esser, K. M., and M. J. Schoenbechler.** 1985. Expression of two variant surface glycoproteins on individual african trypanosomes during antigen switching. Science **229:**190–193.
16. **Fairlamb, A. H., P. O. Weislogel, J. H. J. Hoeijmakers, and P. Borst.** 1978. Isolation and characterization of kinetoplast DNA from bloodstream form of *Trypanosoma brucei*. J. Cell Biol. **76:**293–309.
17. **Jeffreys, A. J., and R. A. Flavell.** 1977. A physical map of the DNA regions flanking the rabbit β globin gene. Cell **12:**429–439.
18. **Liu, A. Y. C., L. H. T. Van der Ploeg, F. A. M. Rijsewijk, and P. Borst.** 1983. The transposition unit of VSG gene 118 of *Trypanosoma brucei*: presence of repeated elements at its border and absence of promoter associated sequences. J. Mol. Biol. **167:**57–75.
19. **Longacre, S., U. Hibner, A. Raibaud, H. Eisen, T. Baltz, C. Giroud, and D. Blatz.** 1983. DNA rearrangements and antigenic variation in *Trypanosoma equiperdum:* multiple expression linked-sites in independent isolates of trypanosomes expressing the same antigen. Mol. Cell. Biol. **3:**399–409.
20. **Michels, P. A. M., A. Bernards, L. H. T. Van der Ploeg, and P. Borst.** 1982. Characterization of the expression-linked gene copies of variant surface glycoprotein 118 in two independently isolated clones of Trypanosoma brucei. Nucleic Acids Res. **10:**2353–2366.
21. **Michels, P. A. M., A. Y. C. Liu, A. Bernards, P. Sloof, M. M. W. Van der Bijl, A. H. Schinkel, H. H. Menke, P. Borst, G. H. Veeneman, M. C. Tromp, and J. H. Van Boom.** 1983. Activation of the genes of variant surface glycoproteins 117 and 118 in *Trypanosoma brucei*. J. Mol. Biol. **166:**537–556.
22. **Michels, P. A. M., L. H. T. Van der Ploeg, A. Y. C. Liu, and P. Borst.** 1984. The activation and reactivation of an expression-linked gene copy of a variant surface glycoprotein in trypanosoma brucei. EMBO J. **3:**1345–1351.

23. **Miller, E. N., and M. J. Turner.** 1981. Analysis of antigen types appearing in first relapse populations of clones of Trypanosoma brucei. Parasitology. **82**:63–80.
24. **Myler, P. J., J. Allison, N. Agabian, and K. Stuart.** 1984. Antigenic variation in African trypanosomes by gene replacement or activation of alternate telomeres. Cell **39**:203–211.
25. **Nasmyth, K. A.** 1982. Molecular genetics of yeast mating type. Annu. Rev. Genet. **16**:439–500.
26. **Pays, E., S. Van Assel, M. Laurent, M. Darville, T. Vervoort, N. Van Meirvenne, and M. Steinert.** 1983. Gene conversion as a mechanism for antigenic variation in trypanosomes. Cell **34**:371–381.
27. **Pays, E., S. Van Assel, M. Laurent, B. Dero, F. Michiels, P. Kronenberger, G. Matthyssens, N. Van Meirvenne, D. LeRay, and M. Steinert.** 1983. At least two transposed sequences are associated in the expression site of a surface antigen gene in different trypanosome clones. Cell **34**:359–369.
28. **Rice-Ficht, A. C., K. K. Chen, and J. E. Donelson.** 1981. Sequence homologies near the C-termini of the variable surface glycoproteins of *Trypanosoma brucei*. Nature (London) **294**: 53–59.
29. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** (1977) Labelling DNA to high specific activity in vitro by nick-translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
30. **Sanger, F., A. R. Coulsen, B. G. Barrell, A. J. H. Smith, and B. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
31. **Schwartz, D. C., and C. R. Cantor.** 1984. Separation of yeast chromosome-sized DNAs by pulsed field gel electrophoresis. Cell **37**:67–75.
32. **Southern, E.** 1975. Detection of specific sequences among fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
33. **Van der Ploeg, L. H. T., A. Bernards, F. A. M. Rijsewijk, and P. Borst.** 1982. Characterization of the DNA duplication-transposition that controls the expression of two genes for variant surface glycoproteins in *Trypanosoma brucei*. Nucleic Acids Res. **10**:593–609.
34. **Van der Ploeg, L. H. T., A. W. C. A. Cornelissen, P. A. M. Michels, and P. Borst.** 1984. Chromosome rearrangements in *Trypanosoma brucei*. Cell **39**:213–221.
35. **Van der Ploeg, L. H. T., D. C. Schwartz, C. R. Cantor, and P. Borst.** 1984. Antigenic variation in *Trypanosoma brucei* analyzed by electrophoretic separation of chromosome-sized DNA molecules. Cell **37**:77–84.
36. **Van der Ploeg, L. H. T., C. Shea, R. Polvere, and M. Gwo-Shu Lee.** 1986. Chromosomal rearrangements and transcription regulation in *Trypanosoma brucei*. *In* N. Agabian, H. Goodman, and N. Noguiera (ed.), Molecular strategies of parasite invasion: UCLA symposia on molecular and cellular biology, new series, vol. 42.
37. **Van der Ploeg, L. H. T., D. Valerio, T. De Lange, A. Bernards, and P. Borst.** 1982. An analysis of cosmid clones of nuclear DNA from Trypanosoma brucei shows that the genes for variant surface glycoproteins are clustered in the genome. Nucleic Acids Res. **10**:593–609.
38. **Vickerman, K.** 1969. On the surface coat and flagellar adhesion in trypanosomes. J. Cell. Sci. **5**:163–193.
39. **Vickerman, K.** 1978. Antigenic variation in trypanosome. Nature (London) **273**:613–617.
40. **Vickerman, K., and J. D. Barry.** 1982. African trypanosomiasis, p. 204–260. *In* S. Cohen and K. S. Warren (ed.), Immunology of parasitic infections. Blackwell Scientific, London.
41. **Williams, R. O., J. R. Young, and P. A. O. Majiwa.** 1979. Genomic rearrangements correlated with antigenic variation in *Trypanosoma brucei*. Nature (London) **282**:847–849.
42. **Young, J. R., J. E. Donelson, P. A. O. Maijwa, S. Z. Shapiro, and R. O. Williams.** 1982. Analysis of genomic rearrangements associated with two variable antigen genes of *Trypanosoma brucei*. Nucleic Acids Res. **10**:803–819.