The IsTat 1.3 VSG multigene family in *Trypanosoma brucei*: retention of the expression linked copy through multiple antigenic switches

Janette Allison, Victoria Rothwell, George Newport¹*, Nina Agabian* and Kenneth Stuart

Issaquah Health Research Institute, 1595 N.W. Gilman Blvd., Issaquah, WA 98027, and

¹Department of Biochemistry, University of Washington, Seattle, WA 98195, USA

Received 1 May 1984; Revised and Accepted 8 November 1984

ABSTRACT

In the IsTar 1 serodeme of *T. brucei* the 3 variant surface glycoprotein (VSG) gene family contains about 10 members, one of which has a telomeric location on a minichromosome. The expression linked copy (ELC) of the 3 VSG gene which occurs in an antigenic variant expressing the 3 VSG, also has a telomeric location but unlike the minichromosomal 3 VSG gene has restriction sites upstream from the 5' barren region. This ELC is retained on the same telomere in (i) a subsequent variant that expresses a telomeric 7 VSG ELC and in (ii) relapse variants and procyclic forms derived from variant antigenic types (VATs) 3 and 7. The 7 ELC has a restriction map upstream from the 5' barren region that differs from, but is similar to, that of the 3 ELC. These data indicate that the 3 and 7 ELCs are on different telomeres when expressed.

INTRODUCTION

A substantial fraction of the African trypanosome genome is devoted to variant surface glycoprotein (VSG) genes. Each cell contains 100-1000 VSG genes which are present as families of sequences varying from one to over ten members (1). During the mammalian phase of the life cycle each trypanosome cell expresses only one of these VSG genes at a time. At a frequency of about 10^{-4} to 10^{-5} per cell division (2) the trypanosomes undergo an antigenic switch which results from the inactivation of the expressed gene concomitant with the activation of another VSG gene. This process occurs repeatedly resulting in the antigenic variation that permits the parasite to evade the host immune response and which is the basis for the remission/relapse characteristic of African sleeping sickness. Upon transfer to the insect vector or appropriate culture conditions, expression of all VSG genes ceases until the mammalian-infective metacyclic stage is produced in the tsetse fly salivary gland (see refs 3-7 for reviews on antigenic variation).

The investigation of the molecular basis for antigenic variation in several laboratories led to the recognition of several key elements. Inactive VSG genes can be located intrachromosomally or adjacent to telomeres, but all

active VSG genes studied thus far are telomeric (see 4, 5, 6). Expression of some VSG genes is associated with the duplication of an intrachromosomal or telomeric basic copy (BC) VSG gene. This copy is inserted into a telomeric site to produce an ELC (8, 9, 10). Expression of other VSG genes is not associated with VSG gene duplication (non-duplication activated, NDA)(11). All genes activated by the NDA mechanism are already located at a telomere. Some telomeric VSG genes can use either the ELC or NDA mechanisms of expression. (12). Recent evidence suggests that some VSG genes which are located on small (100-200 kb) chromosomes may themselves be duplicated for expression (13, 14).

Of particular interest is how only one of the numerous VSG genes is expressed at a time during the mammalian phase of the life cycle and how expression of all VSG genes is inactivated upon transfer to the insect host until the metacyclic stage. Borst and his colleagues (4) have proposed a single (telomeric) expression site model to account for the expression of one VSG gene at the exclusion of all others. While attractive, this model is not supported by studies in which the BoTat 1 VSG gene was mapped in several independently occurring expressors of this predominant variant antigenic type (VAT) of T. equiperdum (10).

In a previous report(15) we described the isolation of VSG gene cDNA clones from several IsTat 1 VATs and presented a preliminary characterization of the 3 VSG gene family in these VATs. These results showed that VAT 3, which expresses the 3 VSG antigen, contains an additional copy (ELC) of the 3 VSG gene. We have further analyzed the 3 VSG gene family and report here that this family is the largest one thus far described for T. brucei. The 3 ELC is located on a telomere in the 3 VSG expressor and can be retained on the same telomere in the genome of subsequent VATs, even if those VATs express their VSG gene by a duplicative mechanism. The presence of the inactive 3 ELC together with an active ELC for another VSG gene suggests that these two genes reside on different telomeres when expressed. Restriction mapping studies reveal that these different telomeres, are related although distinct, thus demonstrating that different VSG genes can be expressed at different genomic sites.

METHODS

Trypanosomes: The production of the IsTat 1 VATs of T.b. brucei has been described previously (16, 17). All VAT populations used in this study were at least 99.9% homogeneous as judged by immunofluorescence using VAT specific

monoclonal antibodies (18). Generation and analysis of the relapse populations expressing the A VSG is the subject of another communication (12). An outline of the lineage of the variants used in this study is given below. While the full designation of these VATs is IsTat 1.X, where X refers to the expressed VSG antigen, the "IsTat 1." is omitted here for convenience. A superscript is used to designate the previous VAT when known. A single trypanosome was used to initiate an acute infection in a rat and the VATs were cloned in lethally irradiated mice from the first parasitemia and first relapse populations on the days indicated. Three hundred trypanosomes from the first relapse population were used to initiate a chronic infection in a Peromyscus mouse from which VATs were cloned on the days indicated. The relapse VATs were isolated from first relapse populations or in first parasitemias in immunized mice and thus probably result from a single antigenic switch as does the origin of VAT 1 from VAT A. The precise lineage and number of antigenic switches leading to VATs 3, 5, 7, and 11, is unknown.

VAT LINEAGE

cdNA Cloning and Preparation of Probes: The production of and preliminary analysis with VSG gene cDNA clones from each of the early IsTat VATs has been reported (15). The cDNA probes were prepared by nick-translation (19) of specific restriction fragments indicated in Fig. 1 that were isolated from a 0.7% (w/v) agarose gel by electroelution (20).

Isolation, restriction digestion and Southern analysis of genomic DNA from trypanosomes: DNA was isolated as previously described (17) and digested with a five-fold excess of restriction enzyme according to the manufacturers specifications. Samples of DNA (3 µg) were fractionated by agarose gel electrophoresis in TBE buffer (0.9 M Tris-HCl, pH 8.35, 0.9 M boric acid, 1.25 mM Na₂EDTA), in 0.5% (w/v) agarose gels for 16 hr at 60 V/cm².

However, mapping studies of the 3 VSG ELC and the "t" genes were performed by electrophoresis of DNA in 0.4% (w/v), 24 cm length, agarose gels for 48-72 hr at 30 V/cm² in which case DNA fragments smaller than 4 kb ran off the gel. After transfer to nitrocellulose membranes (21) the blots were hybridized as previously described in detail (15) in the presence of 10% (w/v) dextran sulphate and washed twice for 30 min. in 3x SSPE, (SSPE is 0.18 M NaCl, 0.01 M Na₂H₂PO₄, 1 mM Na₂EDTA, pH 7.4) 0.25% (w/v) Sarkosyl, at 65°C, then twice for 30 min in 0.1x SSPE, 0.25% (w/v) sarkosyl.

Size fractionation of minimally sheared DNA:

Gel electrophoresis: Live trypanosomes from variants A, 3, and 7 were lysed in the wells of an agarose gel and the released DNA was treated exactly as reported by Sloof *et al.* (22).

Sucrose gradient density centrifugation: To prepare minimally sheared DNA 2 x 10⁹ live cells from VAT 1 were treated according to the procedure of Williams *et al.* (13), layered onto a 10-40% (w/v) sucrose gradient and then centrifuged in a Beckman SW27 rotor at 20,000 rpm for 16 hr at 20°C. The gradient was fractionated through a widebore capillary from the bottom of the tube into 36 x 1 ml fractions. The fractions were processed as described by Williams *et al.* (13) and the resulting DNA was dissolved in 50 µl of 10 mM Tris -HCl, pH 7.5, 1 mM Na₂EDTA. Ten µl of alternate fractions were used to determine the position of 3 VSG genes in the gradient by Bam HI digestion and Southern analysis.

RESULTS:

The 3 VSG multigene family

A detailed restriction map of the 3 VSG structural gene (Fig. 1) was prepared using three plasmids containing overlapping cDNA inserts (see ref.15) One cDNA (pTb1.3-c3) contains the 5' end of the structural gene since it hybridizes with a synthetic 22-mer that is complementary to the 5' terminal 35 nucleotide sequence of VSG mRNAs (23,24,25). Based on the size of other VSG mRNAs, the 1560 nucleotide cDNA sequence represents most, if not all, of the 3 structural gene sequence.

Southern hybridization analyses of the VAT 3 genomic DNA using probes that are derived from the 5' or 3' ends of the 3 VSG cDNA (see Fig. 1) were performed and revealed a large gene family. This is demonstrated in Fig. 2 where DNA from six sequentially isolated variants (A, 1, 3, 5, 7, and 11) was digested with Hind III, an enzyme which cuts the cDNA, and probed with the 5' or 3' specific probe. In general, restriction digests with 16 other enzymes

Figure 1 Restriction map of the 3 VSG structural gene derived from three plasmid cDNA inserts: The probes used in Southern blot experiments described in the text were derived by PstI digestion of pTb1.3-c1 (5' probe) and Bam HI /PstI double digestion of pTb1.3-c2 (3' probe). A, Ava I; B, Bam HI; E, Eco RI; H, Hind III; F1, Hinf I; K, Kpn I; Ps, Pst I; Pv, Pvu II; solid circles indicate Pst I sites regenerated during the cloning procedure.

revealed a minimum of six and a maximum of ten gene copies in all variants analyzed with either the 5' or 3' cDNA probe. However, this variation in gene copy number may be explained by the co-migration of restriction fragments and by restriction site polymorphisms among genes. We estimate that there are ten 3 VSG genes including the ELC. Of these ten copies, two have been isolated from lambda 1059 genomic libraries (asterisks in Fig. 2A) but displayed restriction maps that differed from the map of the expressed gene (data not shown).

A third 3 VSG gene was identified because of its location on a fragment that showed size variations among VATs ("t" in Fig. 2B). This copy appears as a 9.0 kb DNA fragment in VAT A (lane 1), increases in size through the subsequent VATs but then appears as a 4.4 kb DNA fragment in VAT 11 (lane 6). A further example of this size variation can be seen in Fig. 5B where the "t" gene underwent deletions from a 9.5 kb DNA fragment to a 4.6 kb fragment when variant 3 (lane 1) relapsed to variant A³ (lane 2). This variation in DNA fragment size was always detected when restriction enzymes that cut the cDNA were used and the blots were probed with the 3' cDNA probe. This 3' sequence size variation among VATs suggests growth and deletion at the telomere as described by Bernards *et al.* (26). The telomeric location of the "t" gene was confirmed by the presence of a universal restriction site 3' to the gene (15) and by its preferential sensitivity to Bal 31 exonuclease (data not shown). By performing single and double digests with a total of 16 restriction enzymes and probing with the 3' probe, it was possible to approximate a map of the "t" gene in all variants. Resolution of higher molecular weight DNA

Figure 2 Southern blot analysis of the 3 VSG genes in early VATs: DNA isolated from VATs A, 1, 3, 5, 7, and 11 (lanes 1-6, respectively) was digested with Hind III and electrophoresed on a 0.5% (w/v) agarose gel at 60 V/cm² for 16 hr (Panels A and B). DNA from VATs A and 3^{A7} (lanes 1 and 2 respectively) was digested with Pru II and Bam HI and treated similarly (Panel C). After bidirectional Southern transfer, the blots were probed with the 5' probe, or the 3' probe as indicated. Arrows indicate the 3 ELC and lingering ELC in VATs 3, 7 and 3^{A7}, respectively. The asterisk indicates the new ELC in VAT 3^{A7}. The "t" locates the 3 telomeric gene which increases in size until VAT 11 where it undergoes 3' deletions. Stars indicate genes cloned in lambda 1059. Molecular weight markers were derived from a Hind III digest of phage lambda DNA.

fragments was made possible by electrophoresis of the DNA in long gels over several days as detailed in the Methods section. The "t" gene contained all the restriction sites present in the cDNA except it lacked a Kpn I site located at the 3' extremity of the cDNA sequence (see Fig. 1). The absence of the Kpn I site in the "t" gene and its presence in the 3 ELC (see later) allowed us to distinguish these two gene copies. No restriction sites were found 3' to the cDNA hybridizing region of the "t" gene in any VAT DNA. Variation in the "t" gene DNA fragment size among VATs resulted from length changes in the 3' flanking (telomeric) sequences. These sequences varied between 8.3 kb and 8.8 kb through VATs A, 1, 3, 5, and 7 but were represented by a much smaller telomere (3.75 kb) in VAT 11.

Figure 3 Localization of the telomeric 3 VSG gene to a minichromosome: Panels A and B:- DNA from 3 different VATs [A (lane 1), 3 (lane 2), and 7 (lane 3)] was electrophoresed on a 0.5% (w/v) agarose gel after lysis and proteinase K treatment of live trypanosomes in the gel slot. A:- ethidium bromide staining of DNA. B:- autoradiograph of the gel in "A" after hybridization with the 3 VSG 3' probe. Hash marks locate two DNA fractions that migrated slower than phage lambda DNA. Panel C:- DNA isolated from VAT 1 was fractionated a 10-40% (w/v) sucrose gradient. Alternate fractions of this gradient were digested with Bam HI and the hybridization pattern with the 3' probe was determined as described in the legend to Fig. 2. The arrow indicates the position of the telomeric 3 VSG gene. Lane c contains unfractionated DNA.

The fourth copy of the 3 VSG multigene family that we could identify was represented by the 3 VSG ELC. This copy, indicated with an arrow in Fig. 2 (lanes 3), also demonstrated 3' fragment size variation among VATs and Bal 31 exonuclease sensitivity typical of telomeric genes (results not shown). Interestingly, this ELC was retained in the later VAT 7 (Fig. 2 lanes 5, arrows) as described in detail below. The 3 ELC was absent in VATs 5 and 11 which suggests they were not derived from VAT 3 or they lost the ELC as is often observed (4).

The 3 Minichromosome

As mentioned above, the variable DNA fragment of the "t" gene was always detected in Southern analysis with the 3' probe when restriction enzymes that cut the cDNA were used. However, when enzymes that did not cut the cDNA (nor 3' to the "t" gene) were tried, the "t" gene DNA fragment was never observed. This was so even for the VAT 11 DNA in which the "t" telomere is short thus enabling us to detect any restriction sites present within 8-10 kb upstream from the "t" gene. Instead, digests using enzymes that did not cut the 3 VSG cDNA occasionally revealed a large (>50 kb) DNA fragment dependent upon the

Figure 4 A comparative analysis of the restriction maps for the 3 minichromosome, ELC, and cDNA: Restriction sites for the 3 minichromosome and ELC maps were determined as detailed in the Methods section. Restriction endonuclease abbreviations as for Fig. 1 and Bg 2, Bgl II. Values in brackets (< >) indicate stretches of DNA in kb that, for convenience, are not included in the map. The "?" signifies the lack of any measurable restriction sites for over 40 kb 5' to the hybridizing region.

quality of the DNA preparation used (results not shown). One explanation for this result is that the 3 VSG "t" gene is located on a small chromosome (minichromosome) which contains few, if any, restriction sites 5' to the cDNA hybridizing region. Such a chromosome has been described by Williams *et al.* (13) and is only detected in minimally sheared DNA preparations.

To test for the presence of a 3 minichromosome, Southern hybridization analyses using minimally sheared DNA from three variants were performed. Electrophoresis of undigested DNA derived by lysing trypanosomes in the gel slot to prevent shearing as described by Sloof *et al.* (22) resulted in the resolution of two major fractions (Fig. 3A); a higher molecular weight fraction that barely entered the gel and a lower molecular weight fraction (>50 kb). The latter fraction contained a predominant band and one or possibly two larger but less prominent bands. When these fractions were probed with the 3 VSG cDNA both the higher molecular weight fraction and the predominant lower molecular weight fraction were labelled (Fig. 3B) thus substantiating the presence of a 3 VSG gene on a small chromosome in each of the variants tested.

Additional evidence that the 3 gene is on a minichromosome is shown in Fig. 3C. In this instance, minimally sheared DNA from VAT 1 was fractionated by sucrose gradient centrifugation as described by Williams *et al.* (13) and DNA from alternate fractions of the gradient was digested with Bam HI and examined by Southern analysis using the 3' probe. The 9 kb fragment containing

the 3 VSG "t" gene (Fig. 3C arrow) was enriched in the portion of the gradient containing smaller DNA molecules, thus this fragment appeared to derive from a minichromosome. DNA from fractions 14 and 15 was used to map the 3 "t" gene in detail. This map, given in Fig. 4, is compared with the maps of the 3 cDNA and the 3 ELC. The "t" gene was identical to the cDNA when tested with 16 restriction enzymes except for the Kpn I site at the 3' end of the coding sequence. No restriction sites were found 3' to the cDNA-hybridizing region in either the "t" gene or the ELC. A Bgl II site which is also present in the 3 ELC was located 5' to the cDNA-hybridizing region of the "t" gene. Recently, other researchers have reported minichromosomal basic copy sequences (13, 14). Because of its near identity with the cDNA the 3 minichromosome may have potential as a basic copy template in the generation of the 3 ELC. Further support for this notion is provided by the structural similarity of the sequences upstream from the 3 ELC and the "t" gene, that is, the common Bgl II site and the lack of any other restriction sites further 5' to the gene. Thus, although we have been unable to map all 3 VSG gene family members, the minichromosomal 3 VSG gene is a candidate basic copy gene for the ELC that is expressed in VAT 3.

Finally, the 3 minichromosome appears to be relatively stable and has been identified in all VATs examined (13 in total) and in procyclic forms derived from six different variants.

The 3 ELC

The 3 ELC is retained but not expressed in procyclic forms and variants derived directly or indirectly from VAT 3. It was retained in procyclic forms derived from VAT 3 as we have previously shown (27) (Fig. 5A lanes 3 and 4-arrows), and in procyclic forms derived from VAT 7 (Fig. 5A lanes 5 and 6-arrows), which itself possesses the 3 lingering ELC, as indicated earlier (Fig. 2 lanes 5-arrows). The 3 ELC was distinguished from the "t" gene by the presence of the Kpn I site which the "t" gene lacks as described above. The 3 ELC and "t" genes were distinguished from other 3 VSG gene family members by variation in the length of their telomeres among different VATs (see Fig. 2C). The 3 ELC was also retained in later variants whose own VSG genes were activated by either an ELC or NDA mechanism. Firstly, it was present in VAT 7 (Fig. 2 lanes 5-arrows) which is an ELC expressor of the 7 gene (15). However, the 3 lingering ELC in VAT 7 lacks the preferential DNase I sensitivity of the expressed ELC in VAT 3 (results not shown). Secondly, variants 3 and 7 relapsed to variants that expressed an A VSG gene by an NDA (VAT A³) or an ELC (VAT A⁷) mechanism (ref. 12); in both cases the 3 ELC was retained (Fig. 5B compare lanes

Figure 5 Retention of the 3 ELC in procyclic forms and subsequent VATs: Panel A: DNA from the bloodstream (lanes 1, 3, and 5) and procyclic (lanes 2, 4, and 6) forms of VAT 1 (lanes 1,2), VAT 3 (lanes 3,4), and VAT 7 (lanes 5,6) was digested with Bam HI and processed as described in the legend to Fig. 2 using the 3' hybridization probe. Panel B:- DNA from variants 3 (lane 1) and its relapse A^3 (lane 2), 5 (lane 3), 7 (lane 4) and its relapse A^7 (lane 5) and 11 (lane 6) was digested with Hind III and hybridized with the 3' probe. The arrows indicate the 3 ELC and lingering ELC; the "t" locates the 3 telomeric gene.

1 & 2 and 4 & 5- arrows). Finally, when variants A^3 and A^7 themselves relapsed to new variants 5^{A3} and 3^{A7} , respectively, the 3 lingering ELC was still present (results not shown). Although VAT 3^{A7} expressed a 3 VSG, it did this from a newly created 3 ELC and the 3 ELC that originated in VAT 3 remained inactive. This newly created 3 ELC is present along with the 3 L-ELC (Fig. 2C). Thus, the 3 ELC was retained through at least three antigenic switch events ($3 \rightarrow 7 \rightarrow A^7 \rightarrow 3^{A7}$).

The 3 lingering ELC was not present in VATs 5 or 11 (Fig. 5B lanes 3 and 6) which occurred later than VAT 3 in the original infection. Either the 3 ELC was lost in these VATs, or they were not in a cell lineage derived from VAT 3.

Restriction maps were determined for the 3 ELC and lingering 3 ELCs in VATs 3, 7, A^3 , A^7 , 5^{A3} , 3^{A7} and in procyclics derived from VATs 3 and 7. In each case these maps were identical for up to 35 kb 5' to the coding region and

Figure 6 Restriction maps of the 3 ELC in VAT 3, and the 7 ELC in VAT 7: Restriction sites were determined by Southern hybridization analysis as detailed in the Methods section. Only a few of the coding sequence restriction sites are shown. Restriction endonuclease abbreviations as in Fig. 1 legend and Bg 1, Bgl I; Bg 2, Bgl II; Rv, Eco Rv; S, Sal I; Sp, Sph I; St, Sst I; X, Xho I.

varied only in the length of the 3' telomeric sequences by about 0.5 to 1.0 kb. This map is shown in Fig. 6 and displays the characteristic 5' and 3' barren regions that flank telomeric VSG genes (4). Of 16 restriction enzymes tested, only a Bgl II site was found 5' to the coding sequence before the barren region. The data indicate that the 3 ELC present in VAT 3 is retained on the same telomere in all the variants analyzed.

In order to investigate the relationship between the lingering 3 ELC telomere and co-existing telomeres that carry expressed VSG genes we have constructed a restriction map of the 7 VSG ELC. A previous study (15) and unpublished results from our laboratories indicate that the 7 ELC is only present in the 7 VSG expressor. This ELC is generated from one of the two intrachromosomal copies of the 7 VSG gene sequence. Fig. 6 shows that the 3 and 7 ELC maps differed in and around the coding sequence as expected. In addition, while restriction site differences occurred between these two maps 5' to the barren region, the 3 and 7 ELC maps were similar. Of the sites that could be mapped 5' to the barren region in both ELCs, Kpn I, Hind III and Eco RI had different locations in the two maps. The Bgl I, Sph I, Eco RV, Sst I, Xho I and Bam HI sites were present at the same positions in both maps within the limits of experimental error. The 5' barren region was 13 kb in the 3 ELC (between the Bgl II and Pst I sites) and 8 kb in the 7 ELC (between the Sal I and Bgl II sites). While related, these maps have clear restriction site differences.

DISCUSSION

The characteristic features of the 3 VSG gene family in the IsTaR 1 serodeme include, the large size of this family, the presence of one of the genes on a minichromosome, and the retention of the 3 ELC through several VAT switches. The only larger gene family size reported is that of the 28 VSG gene in the BoTaR serodeme of *T. equiperdum* which contains 12 or more copies (10). Of about 10 genes that comprise the 3 gene family, we have characterized four; two from genomic clones, one by its telomeric location on a minichromosome and the fourth by its expression linked duplication to a telomeric site. The large size of the 3 gene family may signify that it is an old gene family or may reflect the ability of this gene family sequence to undergo nucleotide changes to encode several immunologically distinct surface antigens while still retaining significant nucleotide sequence homology to the cDNA.

The telomeric location of both the 3 ELC and the minichromosomal "t" gene was verified by the presence of an apparent universal restriction site 3' to the gene, the variation in size of the 3' flanking sequences and Bal 31 exonuclease sensitivity. Dramatic deletion events were observed in the flanking sequences 3' to the 3 minichromosome gene. This 3' barren region was reduced to about 3.75 kb on two separate occasions. Similar results have been reported by Bernards *et al.* (26) for the 221a gene in which large deletions reduced the 3' barren region to about 3-4 kb. These researchers suggest that the deletions function to check telomeric growth. Apparently, a 3-4 kb 3' segment represents a minimum size for downstream flanking sequences. A mechanism must operate that prevents further deletions 5' into the gene and this may be represented by specific sequences or sequence modifications in the 3-4 kb DNA segment that flanks the gene.

Other investigators (13, 14, 28) have also found VSG genes on minichromosomes. In the IlTat 1.3 (13, 28) and ETat 1.8 (14) VATs, the small chromosomes serve as basic copies and this is a possible function of the 3 minichromosome here. Both the IlTat 1.3 and the 3 minichromosome described here have barren regions for at least 50 to 80 kb 5' to the cDNA hybridizing sequences, and these barren regions may extend to the end of the minichromosomes. Minichromosomes may function to amplify the number of telomeres available to VSG genes. Since telomeric gene conversions can create new VSG gene sequences (29, 30) a larger number of telomeric VSG genes may enhance diversification of the VSG gene repertoire. In addition, telomeric VSG genes may be activated in preference to intrachromosomal VSG genes to produce an antigenic switch. Many early VATs, whether ELC or NDA expressors have their basic copy or NDA gene located near a telomere. These variants include IsTat A, 1, 3,

5,(15), BoTat 1 (31), MiTat 1.2(221a) (4), IlTat 1.2, 1.3, 1.4 (28, 32), ETat 1.8 (14), AnTat 1.1, and 1.3 (9, 33). Furthermore, since VSG genes are only expressed on telomeres, numerous telomeric VSG genes may increase the variety of VATs that are preferentially produced by a particular VAT. Thus, the development of minichromosomes in trypanosomes may have served to create dynamic chromosomal sites that enhance antigenic diversification and influence the sequence of VAT expression.

The retention of the 3 ELC on the same telomere in procyclic forms, in VATs not expressing the 3 VSG and through several VAT switches, demonstrated that the presence of an ELC, even in an expressible genomic environment, does not guarantee its expression. This result is consistent with our previous data (27) which showed that ELCs are retained in procyclics although not expressed. Thus, the property that activated the 3 VSG gene in VAT 3 was not retained by the lingering ELC. A conclusion of these observations is that VSG genes which reside on expressible telomeres (NDA genes) may represent lingering ELCs. This is possible in the case of IlTat 1.3, an NDA expressor, (28) which may have reactivated a lingering ELC, whilst the serologically identical variant, ETat 1.8 (14), created a new ELC for expression. Thus, NDA genes may in many, if not all, cases represent reactivated lingering ELCs.

The lack of the 3 ELC in VATs 5 and 11 which were isolated later in the same infection suggests that either the 3 ELC was lost in these VATs or that these VATs are not in a cell lineage that follows VAT 3. We and others (see 14) have observed that several new VATs can arise from the same VAT and it is possible that variants 5 and 11 were in a different cell lineage from VAT 3. We can be sure that VAT 5 did not switch to the later VAT 7 because of the presence of the lingering ELC in VAT 7. However, we cannot conclude whether VAT 5 arose from VAT 3 (directly or through intermediate variants) and subsequently lost the 3 ELC or whether it arose through a different VAT lineage.

The switch from 3 to 7 entails three major events: the creation of the 7 ELC, the transcriptional activation of the telomeric segment on which the new 7 ELC is located, and the inactivation of the 3 ELC. The presence of the 3 lingering ELC on a previously expressing telomere in VAT 7 demonstrates the physical presence of the expressed 7 ELC on a different telomere. Thus, the 3 and 7 genes are expressed at two different telomeric sites which implies the presence of multiple sites for the expression of VSG genes. This conclusion is supported by the findings of Longacre *et al.* (10) and Young *et al.* (11) who have already reported the presence of multiple telomeric expression sites.

While we cannot exclude the possibility of telomeric exchange as proposed by Borst and Cross (4) and Pay *et al.* (29), we feel this is unlikely, since the restriction map of the 3 ELC remains unchanged for up to 35 kb 5' to the gene. Implicit in the telomeric exchange model is the presence of a promoter 5' to the exchange point which would entail the creation of a large precursor mRNA. The original candidate for this function was the mini-exon spliced leader although recent evidence suggests quite a different role for this sequence (23, 24). Alternatively, a control sequence 5' to the exchange point such as an enhancer might regulate the VSG gene associated promoter from a large distance. No precedent exists for this notion.

The results presented here are consistent with the concept that VSG genes can be expressed at multiple genomic sites as defined by telomeric segments. The existence of alternate expression sites is supported by the recent results of Borst *et al.* (35) and Rothwell *et al.* (unpublished results) who have observed the expression of VSG genes on different chromosome sized DNA molecules. The restriction map similarity between the telomeric segments 5' to the barren region of the 3 and 7 ELCs may reflect nucleotide sequence homologies that function to effect VSG gene conversion, gene transposition, or requirements for sequences that regulate control of VSG gene expression. The existence of multiple expression sites requires that mechanisms must exist that activate one telomeric segment and perhaps coordinately inactivate a previously active gene. The nature of these mechanisms remains a mystery.

ACKNOWLEDGEMENTS

We wish to thank Dr. R.Aline Jr. for sharing unpublished results. We are especially grateful to John Scholler, Elke Gobright and Joanne West for excellent technical assistance. This work was supported in part by grants from NIH (AI 17375 to KS and AI 17309 to NA), UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and DAMD17-92-C-2016 grants to KS.

Abbreviations: cDNA, complementary DNA; EDTA, Ethylenediamine tetraacetic acid; ELC, expression linked copy; IStat(R), Issaquah Trypanozoon antigenic type (repertoire); kb, kilobase pairs; NDA, non-duplication activated; VAT, variant antigenic type; VSG, variant surface glycoprotein.

*Current address: Naval Biosciences Laboratory, 140 Warren Hall, University of California, Berkeley, CA 94720, USA

REFERENCES

1. Van der Ploeg, L.H.T., Valerio, D., De Lange, T., Bernards, A., Borst, P., and Grosveld, F.G. (1982) Nucl. Acids Res. 10, 5905-5923.
2. Doyle, J.J. (1977) in Immunity to Blood Parasites of Animals and Man (Miller, L.H., Pino, J.A., McKelvey, J., Jr. eds.) pp 31-63, Plenum, New

- York.
3. Parsons, M., Nelson, R., Agabian, N. (1984) *Immunology Today* 5, 43-50.
 4. Borst, P. and Cross, G.A.M. (1982) *Cell* 29, 291-303.
 5. Borst, P. (1983) in *Mobile Genetic Elements* (Shapiro J. A., ed.) pp 621-659. Academic Press. New York.
 6. Borst, P., Bernards, A., Van der Ploeg, L.H.T., Michels, P.A.M., Liu, A.Y.C., De Lange, T., Sloof, P., Veeneman, G.H., Tromp, M.C., and Van Boom, J.H. (1983) in *Genetic Rearrangement* (Chater, K.F., Cullis, C.A., Hopwood, D.A., Johnston, A.A.W.B., Woolhouse, H.W., eds.) pp 207-233, Croom-Helm, London.
 7. Englund, P.T., Hajduk, S.L. and Marini, J.C. (1982) *Ann. Rev. Biochem.* 51, 695-726.
 8. Hoeijmakers, J.H.J., Frasc, A.C.C., Bernards, A., Borst P., Cross, G.A.M. (1980) *Nature* 284, 78-80.
 9. Pays, E., Van Meirvenne, N., LeRay, D. and Steinert, M. (1981) *Proc. Natl. Acad. Sci. U.S.* 78, 2673-2677.
 10. Longacre, S., Hibner, U., Raibaud, A., Eisen, H., Baltz, T., Giroud, C., and Baltz, D. (1983). *Molec. Cell. Biol.* 3, 399-409.
 11. Young, J.R., Donelson, J.E. Majiwa, P.A.O., Shapiro, S.Z., and Williams, R.O. (1982) *Nucl. Acids Res.* 10, 803-819.
 12. Myler, P., Nelson, R.G., Agabian, N., and Stuart, K. (1984). *Nature* 309, 282-284.
 13. Williams, R.O., Young, J.R., and Majiwa, P.A.O. (1982) *Nature* 299, 417-421.
 14. Young, J.R., Miller, E.N., Williams, R.O., and Turner, M.J. (1983) *Nature* 303, 196-198.
 15. Parsons, M., Nelson, R.G., Newport, G., Milhausen, M., Stuart, K., and Agabian, N. (1983) *Mol. Biochem. Parasitol.* 9, 255-269.
 16. Stuart, K., Gobright, E., Jenni, L., Milhausen, M., Thomashow, L., and Agabian, N. (1983) *Parasitology*, in press.
 17. Milhausen, M., Nelson, R.G., Parsons, M., Newport, G., Stuart, K. and Agabian, N. (1983) *Mol. Biochem. Parasitol.* 9, 241-254.
 18. Myler, P., Allen, A., Agabian, N., and Stuart, K. *Inf. Immun.* submitted.
 19. Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251.
 20. Girvitz, S.C., Bacchetti, S., Rainbow, A.J., and Graham, F.L. (1980) *Annal. Biochem.* 107, 492-499.
 21. Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517.
 22. Sloof, P., Menke, H.H., Caspers, M.P.M., and Borst, P. (1983) *Nucl. Acids Res.* 11, 3889-3901.
 23. Nelson, R.G., Parsons, M., Barr, P.J., Stuart, K., Selkirk, M., and Agabian, N. (1983) *Cell* 34, 901-909.
 24. Boothroyd, J.C. and Cross, G.A.M. (1982) *Gene* 20, 281-289.
 25. Van der Ploeg, L.H.T., Liu, A.Y.C., Michels, P.A.M., De Lange, T., Borst, P., Majumder, H.K., Weber, H., Veeneman, G.H., and Van Boom, J. (1982) *Nucl. Acids Res.* 10, 3591-3604.
 26. Bernards, A., Michels, P.A.M., Lincke, C.R., Borst, P. (1983) *Nature* 303, 592-597.
 27. Parsons, M., Nelson, R.G., Stuart, K., and Agabian, N. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 684-688.
 28. Young, J.R., Shah, J.S., Matthyssens, G., and Williams, R.O. (1983) *Cell* 32, 1149-1159.
 29. Pays, E., DeLauw, M-F., Van Assel, S., Laurent, M., Vervoort, T., Van Miervenne, N., and Steinert, M. (1983) *Cell* 35, 721-731.
 30. De Lange, T., Kooter, J.M., Michels, P.A.M., and Borst, P. (1983) *Nucl. Acids Res.* 11, 8149-8165.
 31. Longacre, S., Raibaud, A., Hibner, U., Buck, G., Eisen, H. Baltz, T., Giroud, C., and Baltz, D. (1983) *Mol. Cell. Biol.* 3, 410-414.

32. Williams, R.O., Young, J.R., Majiwa, P.A.O., Doyle, J.J., and Shapiro, S.Z. (1980) Cold Spring Harb. Symp. Quant. Biol. 45, 945-949.
33. Laurent, M., Pays, E., Magnus, E., Van Miervenne, N., Matthyssens, G., Williams, R.O., and Steinert, M. (1983) Nature 302, 263-266.
34. Nelson, R.G., Parsons, M., Selkirk, M., Newport, G., Barr, P., and Agabian, N. (1984) Nature 308, 665-667.
35. Borst, P., Bernards, A., Van der Ploeg, L.H.T., Michels, P.A.M., Liu, A.Y.C., De Lange, T., and Kooter, J.M. (1983) Eur. J. Biochem. 137, 383-389.