Volume 10 Number 3 1982 Nucleic Acids Research

Analysis of genomic rearrangements associated with two variable antigen genes of Trypanosoma brucei

John R.Young, John E.Donelson, Phelix A.O.Majiwa, Stuart Z.Shapiro and Richard O.Williams

International Laboratory for Research on Animal Diseases (I.L.R.A.D.), P.O. Box 30709, Nairobi, Kenya

Received 8 December 1981; Accepted 6 January 1982

ABSTRACT

Some variable surface glycoprotein (VSG) genes of Trypanosoma brucei undergo duplication and transposition when they are expressed. We report here the cloning of cDNAs coding for two VSGs from the ILtar 1 repertoire. Analysis of the genomes of trypanosomes expressing these and other antigens shows that there is no additional copy of the sequences coding for either VSG in expressing clones of trypanosomes, and reveals rearrangements analogous to those previously described for the gene for another VSG from this antigen repertoire. The data indicate that duplication does not accompany the expression of these VSG genes. Transposition to a specific expression site cannot be excluded, but would have to involve either a much larger segment of DNA, or movement to a region of much greater homology with the previous flanking sequences, than is observed for VSG genes that are duplicated when expressed. It is reasoned that the control of expression by coupled duplication and transposition is not sufficient to account for the selection of a single VSG gene for expression.

INTRODUCTION

The phenomenon of antigenic variation in the haemoparasitic African trypanosomes involves the ability of the trypanosome to express one at a time from a large repertoire of antigenically distinct surface antigens (variable surface glycoproteins, VSGs). The parasite switches from one antigen to another at a low frequency. The VSG forms a continuous surface coat which probably prevents access by antibodies to other, non varying, surface determinants. Thus chronic infections of mammalian hosts can be maintained, in the face of immune responses against the VSG, by a sufficient rate of antigenic switching to allow the successive emergence of new parasite populations carrying different surface antigens. This phenomenon has been extensively reviewed (1-4). In more recent reviews the possible molecular mechanisms have been discussed (5-7).

Hoeijmakers and coworkers (9,10) and Pays et al. (11,12) have shown that in Trypanosoma brucei, the expression of some VSG genes is accompanied

0305-1048/82/1003-0803$2.00/0

by duplication and transposition to a different site. Trypanosomes expressing a particular VSG have an extra copy of its gene (expression linked copy, ELC) distinguished by differences in flanking regions. Pays et al. have provided evidence that only the duplicated copy is transcribed (14). The extra copy is lost when the VSG's expression is switched off (9-13). Bernards et al.(15) have shown recently that the duplicated copy is inserted into the expression site by recombination close to or within the extreme 3' end of the VSG coding sequence. This information can be integrated in a model for the mutually exclusive expression of VSGs which switches by a process of duplication coupled with replacement of a pre-existing gene in a uniquely transcribable expression site. This model is sufficient only if the initiation of expression of every VSG gene is accompanied by the appearance of a new copy at an expression site.

Our analysis of the rearrangements of the ILtat 1.2 VSG gene (16,17) was not consistent with this model. No extra copy was found in trypanosomes expressing ILtat 1.2 VSG. Both expressing and non-expressing trypanosomes contained two copies of the gene, both of which underwent rearrangements apparently consisting of insertions and deletions of DNA in a region at or beyond the 3' end of the coding sequence. The unduplicated basic copies of those VSG genes for which extra expression linked copies have been observed do not undergo this kind of rearrangement (9-13,33). The pattern of rearrangements observed with the ILtat 1.2 genes did not show any feature that could be correlated with expression. Transposition was not detected, but could not be ruled out.

In this paper we report the preliminary analysis of the genes of two more VSGs from the ILTar 1 repertoire in expressing and non-expressing trypanosome clones. There are multiple copies of both genes, some of which undergo rearrangements exactly analogous to those of the ILTat 1.2 genes. We do not find any duplications of the genes when comparing expressing with non-expressing trypanosome clones. The implications of this observation for the duplication-transposition model of control of VSG gene expressions are discussed.

EXPERIMENTAL PROCEDURES

Trypanosomes. All trypanosomes used were derived from a single clone (clone A) from Trypanosoma brucei stock 227 (1). Clone B1 was derived from a relapse of clone A in a rabbit and Clone B2 was made from a population of clone B1. All clones expressing VSGs C and D were cloned from the first relapse populations in normal mice infected with clone B1. In this paper

trypanosome clones are named by a letter denoting the VSG expressed, followed by a distinguishing number. This nomenclature is identical with that we have used elsewhere (16,17). The formal nomenclature of the VSGs A,B,C and D are, respectively ILTats 1.1, 1.2, 1.3, and 1.4

Trypanosomes were grown in lethally irradiated (900 rads) rats or mice, except for the relapse infections in normal animals from which new variants were isolated. Populations used for isolation of RNA and DNA were shown by immunofluorescence to be at least 99% homogeneous for expression of a single VSG (18).

cDNA cloning. Polyadenylated RNA was isolated from trypanosome clones as described previously (19). These RNA preparations were used as the template for double-stranded cDNA synthesis by the procedures of Buell et al. (20) and Wickens et al. (21) with minor modifications. A partial size fraction of the cDNA was obtained by passing the product of the second strand reaction through a biogel A15M column (0.7 x 18 cm) in 20 mM NaCl. The sharp excluded peak, containing 70-80% of the trichloroacetic acid precipitable radioactivity was treated with S1 nuclease, phenol extracted, and passed through a biogel A150M column (1.5 x 10 cm) in 100 mM Tris-HCl, 300 mM NaCl, 10 mM EDTA, pH 7.4. Fractions containing the first half of the radioactivity eluting from the column were pooled and the DNA precipitated with 2.5 volumes of ethanol at -20°C overnight. A total of 0.3-1.0 µg of double-stranded cDNA was obtained from 3.0 µg of mRNA template in different experiments.

Homopolymer tails of about 20 dCMP residues were added to the 3' ends of the cDNA as described by Gubbins et al. (22). Plasmid pBR 322 digested with PstI was similarly tailed with dGMP residues. About 0.25 µg of dCMP-tailed cDNA and 1 µg of dGMP-tailed PstI-cut pBR 322 were annealed in 1 ml of 100 mM Tris HCl pH 7.4, 300 mM NaCl, 10 mM EDTA by cooling from 60°C to 30°C over a period of 4 hours. Annealed DNA was transformed into E.coli HB101 as described by Rougeon et al. (23). About 1000 tetracycline resistant colonies were obtained from this amount of cDNA, 80-90% of which contained plasmids with DNA sequences inserted at the PstI site of pBR 322 as judged by sensitivity to ampicillin.

Screening of cDNA clones. VSG messenger RNA was purified from trypanosome clones C1 and D1 by the polysome immunoabsorption procedure described by Shapiro and Young (24). Single stranded cDNA was synthesised on these mRNAs using the first strand reaction conditions of Buell et al. (20), except that the concentrations of dATP and dCTP were reduced to 50 µM, and 100 µCi each of (α-^{32}P) dATP and (α-^{32}P) dCTP were added to the 25 µl

reaction. After the reaction, the mixture was incubated in 0.1 NaOH for 60 mins. at 50°C, neutralised, phenol extracted, and passed through a 0.7 x 18 cm Sephadex G-100 column in 20 mM Tris-HCl, 5 mM EDTA, pH 8.0. The excluded fraction contained the labeled cDNA. The specific activities of the labeled DNA were about 5 x 10^8 cpm/µg.

These cDNAs were used as probes for *in situ* colony hybridisation as described by Grunstein and Hogness (25).

Hybrid Arrested translations. Hybrid arrested translation was carried out as described by Paterson et al. (26). The hybridisation mixture contained 5 µg of recombinant plasmid and 1 µg of total polydenylated RNA. The mRNA translation products were synthesized *in vitro* in the rabbit reticulocyte cell free system (Amersham, UK), and analysed by SDS slab gel electrophoresis on 7.5-15% polyacrylamide gels, all as described by Shapiro and Young (24). Immunoprecipitation of cell free translation products to identify the VSG bands was carried out as described by Shapiro and Young (24).

DNA isolation. Nuclear DNAs from trypanosomes were isolated and kinetoplast DNA subsequently removed from the preparation as described by Laurent et al. (27). Plasmid DNA was isolated as described by Citron et al. (28). Restriction enzyme digestions were carried out under conditions suggested by the manufacturers (Biolabs).

Fragments of the plasmid pcBC1 were purified,after separation by polyacrylamide gel electrophoresis as described by Maxam and Gilbert (29). The fragments used in this work are shown in Figure 2.

Filter hybridisation of genomic restriction fragments. Nuclear DNA (1.5 µg) was digested with at least 20-fold excess of restriction enzymes and DNA fragments were separated by electrophoresis in 0.6 or 0.8% agarose gels in 40 mM tris-Acetate, pH 8.2, 2 mM EDTA. The DNA was partially depurinated by soaking the gel for 10-15 minutes in 0.2 N HCl, denatured and cleaved by soaking in 0.5 M NaOH, 1 M NaCl for 30 minutes, neutralized by soaking for 1 hour in 1 M Tris HCl, pH 7.4, 3 M NaCl, and transferred from the gel to nitro-cellulose filters (Schleicher and Schuell) in 20 x SSC. Plasmids and DNA fragments were labeled with ^{32}P by the nick translation procedure described by Rigby et al. (30). Specific activities of 100-200 x 10^6 cpm/µg of DNA were obtained. Hybridisation was carried out using 10% dextran sulphate, as described by Wahl et al. (31). After hybridisation filters were washed twice in 0.1 x SSC/0.1% SDS for an hour at 65°C, dried and exposed to Kodak XR-5 film at -80°C with intensifying screens.

RESULTS

Identification of cDNA clones containing VSG coding sequences. About one thousand bacterial transformants were obtained which contained plasmids with inserted cDNA sequences prepared from a mixture of poly A^+ RNA from trypanosome clones C1 and D1. The colonies were first screened with combined cDNAs made from the immunopurified VSG mRNAs of trypanosome clones C1 and D1. Thirty colonies which gave strong positive signals were further screened on replicate filters using the separate cDNA probes. Colonies which hybridized strongly to one probe and negligibly with the other were tentatively identified as containing VSG coding sequences. No colonies hybridized significantly to both probes.

Hybrid arrested translation was then used to show that two of these recombinant plasmids, pcBC1 and pcBD1, contained coding sequences for VSGs C and D respectively (figure 1). In each case (lanes 1, 2, 3 for VSG-C and lanes 4, 5, 6 for VSG-D) hybridization of the poly A^+ RNA with the plasmid sequences prior to translation eliminated, or largely removed, a band on

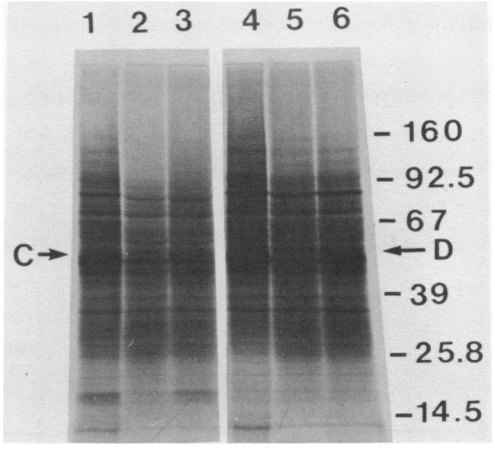

Figure 1: Demonstration by hybrid arrested translation that plasmids pcBC1 and pcBD1 contain coding sequences for VSGs C and D respectively. Tracks 1-3; translation products of mRNA from trypanosome clone C1: track 1 - untreated, track 2 - after hybridisation to pcBC1, track 3 - after hybridisation to pcBC1 then boiling to denature the hybrids. Tracks 4-6; the same series of experiments with mRNA from clone D1 and plasmid pcBD1. The position of marker proteins in kilodaltons is shown at the right. Arrows labelled C and D show the position of in vitro synthesised VSGs C and D identified by immunoprecipitation (not shown). Inhibition of synthesis with the plasmids which is reversed on denaturation demonstrates homology of the inserted sequences with the VSG mRNAs.

the gel corresponding to a protein of about 60,000 daltons. This band reappeared if the RNA:DNA heteroduplex was denatured by boiling before adding to the in vitro translation system (lanes 3 and 6). The band co-migrated with immunoprecipitated VSG in each case (not shown). This result, coupled with the original screening with immunopurified mRNA, provides very strong evidence that the VSG coding sequences had been obtained.

The sequences of the cDNA inserted in these two plasmids have been determined (32). The finding of 3'-terminal sequences homologous to other VSG mRNAs supports the identifications of these plasmids. A map of those restrictions enzyme sites in pcBC1 relevant to the data in this paper is shown in figure 2. Plasmid pcBC1 contains the whole of the 3' end of the VSG C mRNA sequence including part of the poly A tail. Plasmid pcBD1 does not contain the 3' extremity of the VSG D coding sequence. The lengths of the inserted sequences in pcBC1 and pcBD1 are 1620 and 770 base pairs respectively.

VSG C genes. The enzyme Hinc II does not cut the cDNA sequence in pcBC1. Figure 3 shows the hybridisation of all but the extreme 3' 500 base pairs of pcBC1 (fragment 3 of figure 2) to Hinc II fragments of nuclear DNA from seven trypanosome clones, two expressing VSG C (C1 and C4), and five expressing other VSGs. Apart from an extra faint band in clone C1, four hybridising fragments are found in all trypanosome clones. Two of these (4.8 and 1.1 kb) are of identical length in all clones and two have sizes varying between 7.5 and 14 kb in different clones. Comparison of the sizes of the varying fragments of clones C1 and C4 with those in the other clones does not reveal any correlation between the length variation of these

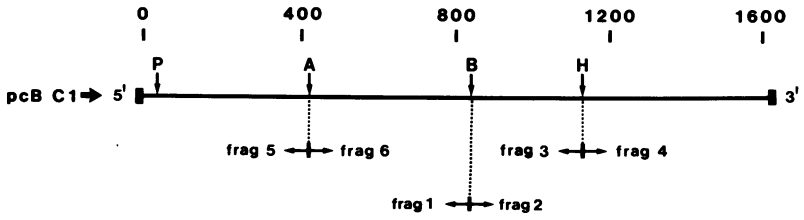

Figure 2: Simple restriction map of the cDNA insert of pcBC1 derived from the published sequence (32). The scale is in base pairs. Bars at the ends represent GC boundaries. Enzyme sites are A: Ava 1, B: Bgl II, H: Hind III, P: PstI. The 5' and 3' ends of the coding sequence deduced from the sequence, are marked. Fragments used in the experiments of figure 3, 4 and 5 span from the enzyme sites shown in the indicated direction into the pBR 322 sequences flanking the cDNA.

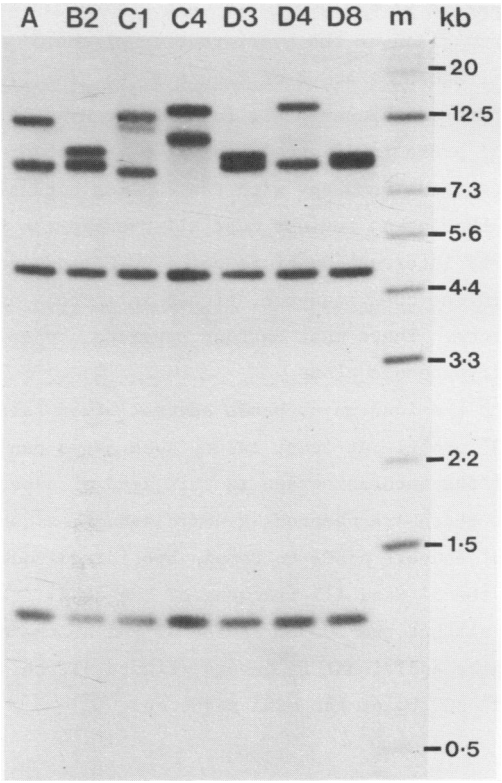

Figure 3: Hybridisation of 5' 1100 base pairs of pcBC1 (fragment 3 or figure 2) to Hinc II fragments of DNA from various trypanosome clones. The clone from which DNA was used is indicated above each track.

fragments and the expression of the gene. There is no apparent relationship between the size changes observed for each fragment in the comparison of any two clones. We believe that the faint band seen in clone C1 is the result of a heterogeneity of the trypanosome population with regard to the rearrangements affecting the varying fragments, there being a corresponding reduction in intensity of the largest band. It is not present in clone C4, and is therefore not a characteristic of the expression of VSG C. This interpretation is supported by the data concerning the VSG D genes which is presented and discussed below.

Detection of four hybridising fragments with an enzyme, Hinc II, that does not cut the probe sequence, suggests that there are four copies of the VSG gene, but could also result from the interruption of one or more of a

smaller number of copies by a Hinc II site not present in the cDNA sequence. Figure 4 shows a comparison of the hybridisation of the 5' and 3' halves of the cDNA sequence (fragment 1 and 2 of figure 2) to Hinc II fragments of DNA from three trypanosome clones. All four of the Hinc II genomic fragments detected with the 5' probe in figure 3 hybridise to both of these probes, showing that each contains homology with the cDNA on both sides of the Bgl II site. Since these homologous regions must all overlap, the four fragments cannot be produced by interruption of genomic VSG C sequences by Hinc II sites not present in the cDNA sequence, either in introns or as the result of sequence divergence. There must be four separate copies of VSG C coding sequences in each trypanosome clone.

In addition to the four major bands several other fainter bands are detected with the 3' probe. At least twenty such bands can be found by longer exposures of the autoradiograph in this kind of experiment (not shown). This indicates that there are numerous genomic sequences containing sufficient homology with the 3' end of pcBC1 to remain hybridized under stringent washing conditions. Since the 5' Hind III fragment of the cDNA (fragment 3 of figure 2) contains all but the 500 pb at the 3' end of the cDNA, and does not hybridise to these additional fragments (figure 3), this homology is restricted to the 3' 500 bp of the mRNA sequence.

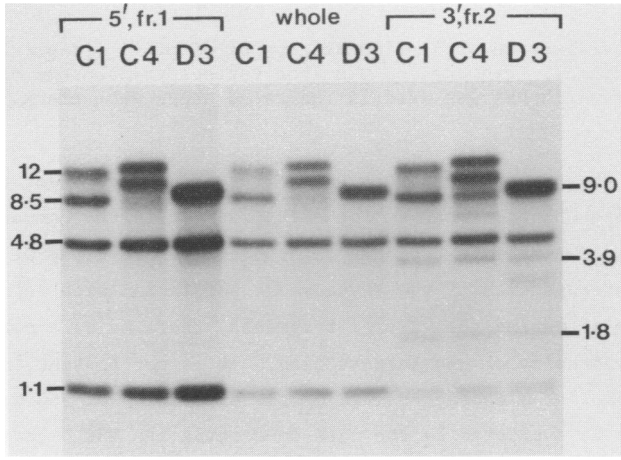

Figure 4: Hybridisation of pcBC1 and fragments from 5' and 3' sides of the Bgl II site (frag 1 and 2 of figure 2) to Hinc II fragments of DNA from three trypanosome clones. Sizes of fragments marked were measured using markers hybridising to pBR 322 as in figure 3 in several different experiments.

If all four copies of the VSG C gene also contain the region spanning the unique Ava 1 site (see fig 2), then we would expect to find eight fragments in Ava 1 digest of trypanosome DNA hybridising to the whole cDNA, of which four would hybridise to each of the 5' and 3' fragments separated by the Ava 1 site in the cDNA (fragments 5 and 6 of figure 2). Hybridisation of these probes to Ava 1 digests of the same DNAs used in figure 4 is shown in figure 5a. With the whole cDNA we find seven bands with each DNA (the 7.3 kb band in clone D3 is a doublet). Three bands 3.3, 4.5 and 12.5 kb, are identified as being 5' end fragments. The 4.5 kb band is more intense than the others, suggesting that two of the four copies have identical 5' Ava 1 fragments. Alternatively one copy may be missing the extreme 5' end of the cDNA sequence. There are four 3' end fragments from each DNA (in addition to faint bands from sequences with limited 3' end homology) as expected. Two

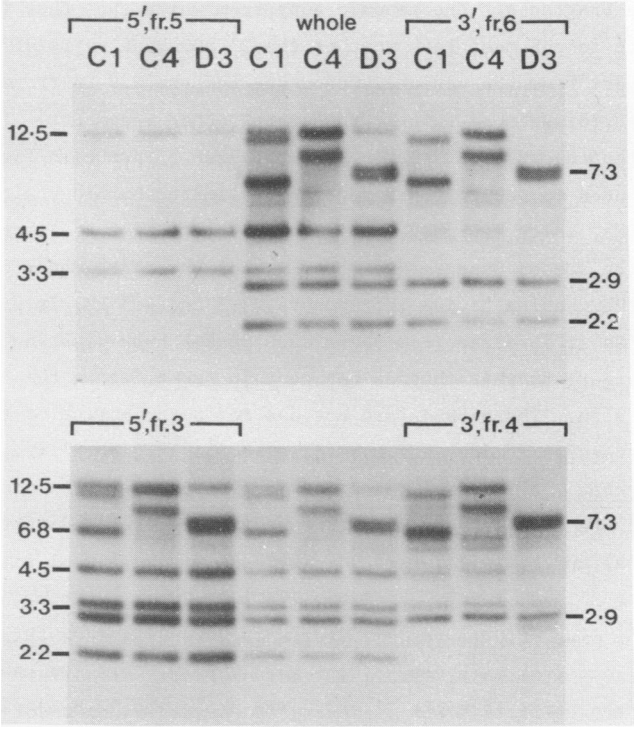

Figure 5: Hybridisation of pcBC1, and 5' and 3' fragments, to DNA from three trypanosome clones digested with Ava 1. The probes are (5a) 5' and 3' Ava 1 fragments of pcBC1 (frags 5 and 6 of figure 2) (5b) 5' and 3' Hind III fragments of pcBC1 (frags 3 and 4 of figure 2). The Central three tracks in each gel were hybridised to the whole of pcBC1.

of the 3' fragments show clone specific length variations whose pattern is identical to that of the varying Hinc II fragments. The same pattern of length variation in 3' fragments is found with six other enzymes (not shown). In no case have we found any variation in 5' fragments.

Hybridisation of the 5' and 3' Hind III cDNA fragments (3 and 4 of figure 2) to the same Ava 1 digests is shown in figure 5b. The long 5' probe used in this experiment hybridises strongly with all four of the 3' Ava 1 fragments detected in figure 5a. It does not hybridise to the large number of genomic sequences with only limited 3' end homology with the cDNA (compare figures 3 and 4). Hybridisation of the variable 3' fragments to this probe shows that the length variation is occurring in fragments which contain a large part of the cDNA sequence, and which are not members of the family of sequences with limited homology. The short 3' probe hybridises to only three of the four 3' Ava 1 fragments detected in figure 5a, including both variable fragments. The genomic copy from which the other (2.2 kb) 3' Ava 1 fragment is cut must lack most of the 3' 500 base pairs of the cDNA sequence. Apart from the variable fragments and the 2.9 kb fragment, the 3' probe also hybridises more faintly to several other fragments. The same bands are detected extremely weakly by the longer 3' probe in figure 5a. They can be attributed to the multiple regions of limited homology with the 3' end of the cDNA, which represents a larger proportion of the shorter probe sequence in the second experiment. That two of these fragments comigrate with two of the genuine 5' Ava 1 fragments (3.3 and 4.5 kb) is probably a coincidence, as it is clear from their much weaker hybridisation to the long 3' probe in figure 5a that they do not contain sequences on the 5' side of the Hind III site. These bands are not due to contamination of the probe with 5' sequences as it does not hybridise to the 12.5 kb 5' fragment.

VSG D genes. A fragment was prepared from pcBD1 containing the 3' 500 base pairs of the cDNA, and hybridised to trypanosome DNAs digested with Hae III (figure 6a) and Mspl (figure 6b). There is a single Mspl site 700 base pairs from the 3' end of pcBD1. Since the probe does not overlap this site it should hybridise only to genomic fragments extending from this site in 3' direction, one from each copy of the gene. pcBD1 has two Hae III sites 440 and 750 base pairs from the 3' end. The probe should hybridise only to a 310 base pair internal fragment and to a fragment running from the more 3' site to beyond the 3' end of the cDNA sequence for each genomic copy. Therefore with both enzymes we expect to find one 3' fragment for each copy of the gene. With both enzymes there are three major bands from the DNA of

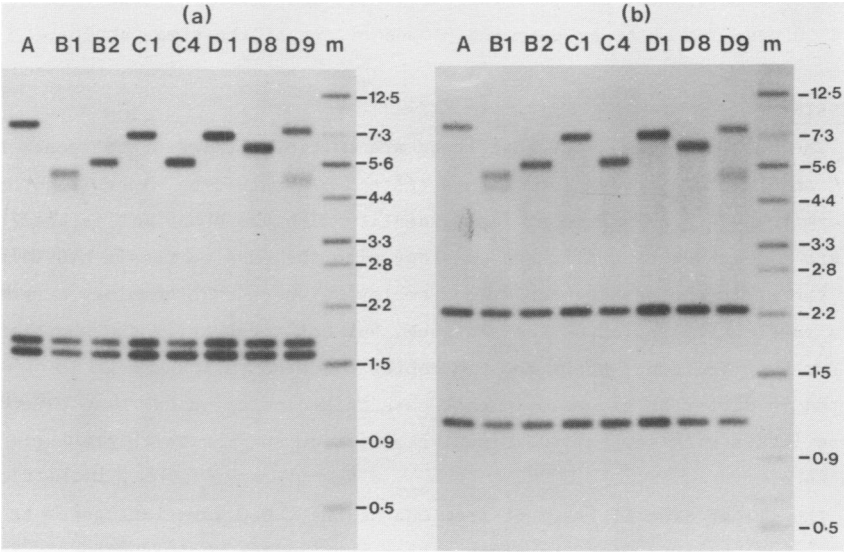

Figure 6: Hybridisation of the 3' 500 base pairs of pcBD1 to DNA from several trypanosome clones digested with Hae III (5a) and Msp 1 (5b).

each trypanosome clone. In both cases two bands are constant and the third varies between different clones. There is no apparent relationship between the size of this band and the expression of the gene. In two of the trypanosome clones, B1, and D9, there is an additional faint band. The appearance of the extra band is clearly not a characteristic of expression as it is not found in other expressors (D1, from which the cDNA was obtained, and D8) and it is found in a non expressor (B1). It probably reflects a heterogeneity within these particular trypanosome populations. This interpretation is supported by the fact that in both of these clones the relative intensity of the larger varying fragment is correspondingly reduced. The broadness of the faint band in clone D9 suggests that there may be a further limited heterogeneity within this population.

As was the case with the varying fragments hybridising to pcBC1, the pattern of size changes seen with these two enzymes is the same. Indeed the fragments are of near identical lengths, with the two enzymes. The failure of this 3' probe to reveal a large family of related sequences hybridising weakly, as observed with the VSG C 3' probes, is not surprising as the VSG D cDNA does not contain the 3' end of the mRNA sequence (32).

These data suggest that there are three copies of the VSG D gene in

all trypanosome clones, and that rearrangements analogous to those seen with the VSG C gene, occur at or near the 3' end of one of the three copies.

DISCUSSION

The data described show that there are multiple copies of the genes for VSGs C and D in all trypanosome clones tested. It should be emphasised that the term "copy" is not meant to imply identity with the cDNA. It is used to indicate a sequence of sufficient homology with the cDNA to remain hybridized under the stringent washing conditions employed, where that homology extends over a substantial region of the sequence, but not necessarily the whole of it. For both VSG genes there are two copies where no rearrangement is detected in different clones. In each case this conclusion has been extended by experiments with several additional restriction enzymes which are not shown here.

The 1.1 kb Hinc II fragment from one of the VSG C invariant genes is shorter than the cDNA, and therefore cannot contain the whole sequence. Its hybridisation to the 3' Bgl II fragment (figure 4) is notably weaker than that of the other three major fragments detected with the 5' probe. This is not so with the 5' probe and shows that this fragment lacks sequences at the 3' end of the cDNA. The observation that one of the VSG C genomic 3' Ava 1 fragments does not hybridise to the 3' Hind III cDNA probe (figure 5b) suggests that this comes from the same gene copy which is lacking most of the 3' 500 base pairs of the cDNA sequence.

With two copies of the VSG C gene and one copy of the VSG D gene there are rearrangements affecting only fragments extending into the 3' flanking sequences. For each of the three fragments concerned the changes in fragment lengths observed are the same with several different enzymes. This suggests that the distal end of these fragments is in a region of identical structure in all trypanosome clones, and that the rearrangements may consist of insertions and deletions of DNA between the 3' end of the gene and this conserved structure, rather than transposition. This is exactly analogous to our previous observations with the ILtat 1.2 gene (16,17) where the same kind of rearrangement was detected at the 3' end of both copies of this gene. The insertions and deletions could be the result of unequal crossing over between repetitive sequences in the 3' flanking regions of these gene copies. Many restriction enzymes with four base recognition sequences (e.g. Hae III and Mspl, figure 6) do not cut between the 3' end of the gene and the conserved distal structures beyond the region of length variation, giving

unusually large fragments for such enzymes. This is suggestive of repetitive sequences. The distal ends of the recombining sequences would have to be homologous to produce the coordinated changes in the sizes of these fragments. Since the changes in length of fragments next to each of these genes are different in comparing pairs of trypanosome clones, we have so far identified five distinct locations in the genome which are subject to this kind of rearrangement (two each for VSGs B and C, one for VSG D).

Duplication of a VSG gene in expressing clones could be revealed in either of two ways; the appearance of an extra band, or the doubling of intensity of a band already present in non-expressing clones. In some clones expressing VSGs C and D (e.g. C1 and D9) we have found additional bands. They do not represent members of the families of genomic sequences with 3' homology, since they are detected with probes that do not hybridise to such sequences (figures 3 and 6). However, they are relatively faint, always accompanied by a decrease in intensity of another variable size band, not present in all expressers (C4, D1, D8) and sometimes found in non-expressers (with VSG D in clone B1). We are confident that they result from a heterogeneity in the trypanosome populations with regard to the 3' length variation described. A similar phenomenon was observed with the ILTat 1.2 gene (16,17). It is not accompanied by heterogeneity of VSG expression.

A duplicated copy of a gene might give rise to fragments of identical size with the fragments from the pre-existing gene from which it was copied, so that no new bands are observed. In this case we would expect an increase in intensity only of those bands which are produced from the gene copy which is duplicated. Different bands derived from the other copies of the genes for VSGs C and D provide an internal control within each gel track, by which any increased intensity can be judged. These intensity increases of bands from VSG genes known to be duplicated, are readily observable (11,33). We have not detected any such intensity changes in the experiments shown here and many others with VSG C and D cDNAs. Thus we do not observe the presence of an extra copy of either VSG C or VSG D gene in expressing clones by either of its possible manifestations.

We have now shown that switching on and off of three ILTar 1 VSGs (1.2; refs. 16 & 17, 1.3 and 1.4; here) does not involve the appearance and subsequent loss of an extra copy of the VSG genes. This is in contrast with the observations made with other VSG genes (9-12), including one within the ILTar 1 repertoire observed in our laboratory (33), where insertion of a duplicated copy into a specific site accompanies the

expression of the gene, and the duplicated copy is lost either upon switching to expression of a different VSG, or, in one case, after transformation and propagation in procyclic culture (11). It appears that there are two different groups of VSG genes. In one group duplication is strictly linked to expression; in the other expression is controlled by a mechanism not involving duplication. Nevertheless it seems unlikely that different VSG genes are regulated by completely different mechanisms.

There are two ways in which we can attempt to reconcile the differences. It could be that one of the multiple copies of some genes, though not duplicated when expressed, is transposed into an expression site. Transposition would be expected to result in changes in flanking sequences on both sides of the gene. Recombination within the 3' end of VSG mRNA sequences has been shown to occur when expression-linked copies of some VSG genes are inserted at the expression site (15, J.E.D. unpublished observations). For two copies of the VSG C gene, fragments extending beyond the 3' end of the cDNA, which includes the 3' end of the mRNA sequence, are identical in expressing and non-expressing trypanosome clones (e.g. figure 5). For the copies whose 3' end fragments vary in length, the coordinated change in sizes of these fragments with many enzymes indicates that there is a structure at the distal end of these fragments which is the same in expressing and non-expressing clones. Therefore, if the kind of 3' end recombination reported for ELC insertion is involved, it must occur between regions sufficiently homologous that both contain this structural feature. The same arguments apply to the 3' fragments of the VSG D genes. We have not detected any changes in 5' flanking sequences of VSG C genes. Despite this evidence however, it is still possible that transposition has occurred, either by movement of a large segment of DNA which encompasses the conserved features at both ends, or by movement of a smaller segment to a region sufficiently similar that it includes those features. The former would require co-transposition of at least 3 kb upstream of the 5' end of the cDNA sequence (the smallest 5' Ava 1 conserved fragment, figure 5), which implies a larger transposed segment than has been observed with VSG genes that are duplicated for expression (10,12). The latter would be in contrast to the different flanking environments detected for duplicated VSG genes (9-13,15).

Another possibility is that one (or more) of the multiple copies of genes for VSGs C and D has been produced at some time by a duplication mechanism similar to that observed for expression-linked copies of other genes, and differs only in that it is maintained in non-expressing clones

both preceding and following its expression. By analogy with the constant environment of basic copy genes (which are duplicated to produce ELCs) (9-12), the non-varying copies of VSGs C and D might be the equivalent of basic copies, and the varying copies the equivalent of ELCs. The analogy is extended by the fact that 3' distal differences between expression-linked copies of one VSG gene in separately isolated clones expressing that VSG, have been observed by Pays et al. (12). This model would not fit the data obtained with the ILTat 1.2 gene (16,17) where both copies show 3' length variation. It would have to be supposed that a pre-existing, basic copy-equivalent had been lost.

If this interpretation is correct then, since potentially expressable copies of genes for VSGs B, C and D co-exist in trypanosome clones expressing only one of them, there must be a mechanism which ensures that only one is selected for expression. These copies also exist in clone A which is expressing a VSG gene whose expression is linked to duplication. Therefore the selection mechanism cannot be independent of that controlling the expression of ELC genes. It has been proposed (15) that the mutually exclusive expression of VSG genes is a result of the availability of only a single functional expression site. If this applies also to the VSG B, C and D genes then the variable copies of these genes must be transposed to this site in a manner consistent with the constraints noted above. The distinction between the two groups of VSG genes would then be that one group can be duplicated only directly into the expression site, whereas the other group can be duplicated, maintained, and then independently transposed into and out of the expression site. This would explain why, for the ELC genes, duplicated copies are found only in expressing clones. With the VSG B, C and D genes we have not detected the postulated transposition. It remains equally possible that some mechanism other than transposition is responsible for the selection among expressable copies of these genes. Whether the 3' changes we have observed might be involved in this mechanism cannot yet be decided. Even if transposition is not involved it is still possible that insertion of an ELC at an expression site (not necessarily unique) might somehow exclude the expression of the other group of genes. Equally the ELC expression site might itself be one of the elements under the control of the same mechanism which selects among those genes for which potentially expressed copies are maintained. These alternatives are amenable to experimental test. They also raise an interesting question concerning the evolution of the mechanism controlling VSG expression. It is possible that

duplication-transposition was the earliest mechanism whereby trypanosomes could switch between expression of different surface antigens, and the maintenance of expressible copies of some genes was a later development requiring a further control mechanism. Alternatively the mechanism of selection without duplication might have preceded the duplication and recombination process whose evolution would then have provided a means of expanding greatly the VSG repertoire by bringing a potentially unlimited number of genes, via the expression site, under the control of the more primitive mechanism.

ACKNOWLEDGMENTS

John E. Donelson was supported in part by a grant from the Fogerty International Center of the U.S. Public Health Service while on leave from the department of Biochemistry, University of Iowa, U.S.A. We thank Dr. J.J. Doyle for trypanosome clones and for carrying out the immunofluorescence, Dr. G. Matthyssens for critical reading of the manuscript, and Miss C. Munyua and Mrs. R. Ndumbu for typing.

REFERENCES

1. Doyle, J.J. (1977) in Immunity to Blood Parasites in Animals and Man, (Miller,L., Pino, J. and McKelvey, eds.) pp. 27-63, Plenum Press. New York, N.Y.
2. Vickerman, K. (1978) Nature 273, 613-617.
3. Cross, G.A.M. (1979) J. Gen. Microbiol. 113, 1-11.
4. Cross, G.A.M. (1978) Proc. R. Soc. B 202, 55-72.
5. Turner, M.J. and Cordingley, J.S. (1981) in Molecular and Cellular aspects of microbial evolution, Carhte, Collins and Moseley (eds.), Cambridge University Press.
6. Englund, P.T., Hajduck, S.J. and Marini, J.C. (1982) Ann. Rev. Biochem. 51, (in press).
7. Marcu, K. and Williams, R.O. (1981) in Genetic Engineering (Setlow, J.K. and Hollaender, E. eds) pp. 129-155 Plenum Press, New York, N.Y.
8. Sakano, H., Huppi, K., Heinrich, G. and Tonegawa, S. (1979) Nature 280, 288-291.
9. Hoeijmakers, J.H.J., Frasch, A.C.C., Bernards, A., Borst, P. and Cross, G.A.M. (1980) Nature 284, 78-80.
10. Borst, P., Frasch, A.C.C., Bernards, A., Hoeijmakers, J.H.J., Van der Ploeg, L.H.T. and Cross, G.A.M. (1980) Am. J. Trop. Med. Hyg. 29, 1033-1036.
11. Pays, E., Van Meirvenne, N., Le Ray, D. and Steinert, M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 2673-2677.
12. Pays, E., Lheureux, M. and Steinert, M. (1981) Nucleic Acids Res. 9, 4225-4238.
13. Agabian, N., Thomashow, L., Milhausen, M. and Stuart, K. (1980) Am. J. Trop. Med. Hyg. 29, 1043-1049.
14. Pays, E., Lheureux, M. and Steinert, M. (1981) Nature 292, 265-267.
15. Bernards, A., Van der Ploeg, A., Frasch, A.C. and Borst, P. (1981) Cell (in press).
16. Williams, R.O., Young, J.R., Majiwa, P.A.O., Doyle, J.J. and Shapiro, S.Z. (1980) Cold Spring Harbor Symp. on Quant. Biol. 45, 945-949.

17. Williams, R.O., Young, J.R., Majiwa, P.A.O., Doyle, J.J. and Shapiro, S.Z. (1980) Am. J. Trop. Med. Hyg. 29, 1037-1042.
18. Doyle, J.J., Behin, R., Mauel, J. and Rowe, D.S. (1975) Anal. N.Y. Acad. Sci. 254, 315-325.
19. Williams, R.O., Marcu, K.B., Young, J.R., Rovis, L. and Williams, S.C. (1978) Nucleic Acids Res. 5, 3171-3182.
20. Buell, G.N., Wickens, M.P., Payrer, F. and Schimke, R.T. (1978) J. Biol. Chem. 253, 2471-2482.
21. Wickens, M.P., Buell, G. and Schimke, R.T. (1978) J. Biol. Chem. 253, 2483-2495.
22. Gubbins, E.J., Maurer, R.A., Hartley, J.L. and Donelson, J.E. (1979) Nucleic Acids Res. 6, 915-930.
23. Rougeon, F., Kauvilskey, F. and Mach, B. (1975) Nucleic Acids Res. 2, 2365-2378.
24. Schapiro, S.Z. and Young, J.R. (1981) J. Biol. Chem. 256, 1495-1498.
25. Grunstein, M. and Hogness, D. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 3961-3965.
26. Paterson, B.M., Roberts, B.E. and Kuff, E.L. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4370-4374.