Volume 10 Number 2 1982 Nucleic Acids Research

Characterization of the DNA duplication-transposition that controls the expression of two genes for variant surface glycoproteins in Trypanosoma brucei

L.H.T.Van der Ploeg, A.Bernards, F.A.M.Rijsewijk and P.Borst

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

Received 26 October 1981; Revised and Accepted 9 December 1981

ABSTRACT

The genome of Trypanosoma brucei carries over a hundred genes coding for different variants of the major surface glycoprotein. Activation of some of these genes is accompanied by a duplication and transposition of the gene (the basic copy) to another region in the genome where it is transcribed. We present here physical maps of the basic and transposition-activated genes for two surface glycoproteins of Trypanosoma brucei, stock 427. In both cases the transposed segment starts 1-2 kb in front of the coding region and ends within the 3'-terminal region of the gene. The DNA segments flanking both transposed genes are indistinguishable and share a 6-kb stretch upstream and a 8-kb stretch downstream of the transposed segment not cut by several restriction endonucleases.

The 5' borders of the two transposed segments are homologous and contain sequences present in many copies in the genome. A different repeated sequence has previously been found at the 3' edge of the transposed segment. The replicative transposition may, therefore, involve a unidirectional gene conversion initiated by base pairing between the edges of the transposed sequence and a single expression site elsewhere in the genome.

INTRODUCTION

Antigenic variation in trypanosomes is the result of the sequential synthesis of different variants of the major surface protein, the variant surface glycoprotein (VSG) [1-3]. Each VSG is specified by a prominent polyadenylated messenger RNA (mRNA) [4-7]. DNA complementary to several of the VSG mRNAs has been cloned as recombinant DNA in Escherichia coli [8-11] and these complementary DNA (cDNA) clones have been used to study VSG genes and the control of their expression [8,9,11-24a].

Our group has isolated and characterized cDNA clones corresponding to the VSG mRNAs from four variants of T. brucei, stock 427, designated 117, 118, 121 and 221 [9]. These cDNAs have been used to study the sequence of translated and untranslated regions in VSG mRNA [14,16,17] and the structure and

0305-1048/82/1002-0593 $2.00/0

organization of VSG genes in genomic blots [11-13,15,18]. The gene analysis has demonstrated that each trypanosome in a stock contains the entire VSG gene repertoire of that stock [11,15]. Comparison of nuclear DNA from trypanosome variants that make different VSGs has shown that activation of the 117 and 118 genes results in the duplication of these genes and the transposition of the extra copy to an expression site [11-13,15,23]. The extra copy is called the expression-linked copy (ELC), the other the basic copy (BC). The ELC is lost when its expression is switched off and it reappears (in altered form) when the gene is switched on again [18,23]. Proof that the ELC is the copy used for mRNA synthesis has come from our recent demonstration that the 117 and 118 BC genes lack the 3'-end present in the corresponding mRNAs [18]. This end is added on to the gene in the duplication-transposition that gives rise to the ELC [16-18]. Evidence compatible with a duplication-transposition of expressed VSG genes has also been obtained by Steinert's group for another stock of T. brucei [23,24a]. In this case, a segment corresponding to the 5' half of the ELC was preferentially degraded when isolated nuclei were digested with DNase I, indicating active transcription of this gene [24].

To further characterize the duplication-transposition that controls VSG gene expression, a more detailed comparison of BC and ELC genes for several variants is required. This comparison has turned out to be less straightforward than one would anticipate. VSG genes exist in families [11-13,15]and a probe for one gene often cross-hybridizes to several related genes, complicating blot analysis of restriction enzyme recognition sites in and around VSG genes in total nuclear DNA. It is also not possible to rely solely on comparison of cloned VSG gene copies, because attempts to clone ELCs have failed thusfar, even though the corresponding BCs are readily cloned in phage lambda [15] and/or cosmid vectors [Grosveld, F.G., Flavell, R.A. and Van der Ploeg, L.H.T., unpublished]. This selection against ELCs in nuclear DNA clone banks may be related to the unusual properties of the DNA regions surrounding the ELCs (see below).

By combining the nuclear blot analysis with an analysis of cloned BC genes we have succeeded in constructing physical maps

of the BCs and ELCs of the 117 and 118 VSG genes. The results of this analysis define the lengths of the transposed segments and show that both genes are transposed into the same (or very similar) expression site(s). In addition, the presence of repeated DNA segments at the borders of the transposed segments provides a hint at the mechanism responsible for the transposition.

MATERIALS AND METHODS

Isolation of trypanosome nuclear DNA. T. brucei, stock 427, was grown in rats and purified free from blood elements as described [25]. Trypanosome DNA was prepared by standard procedures. Briefly, trypanosomes (1 g wet wt.) are suspended in 10 ml cold NET buffer (NET = 100 mM NaCl, 100 mM EDTA, 10 mM Tris-HCl (pH 7.5)) and lysed by the addition of sodium dodecylsulphate to 1%. After lysis proteinase K is added to 50 µg/ml and the suspension is incubated for 30 min at 37°C. After phenol extraction and ethanol precipitation, the DNA is dissolved in 5 ml of 40 mM Tris-HCl (pH 7.5), ribonuclease is added to 20 µg/ml and the digestion is allowed to proceed for 30 min at 37°C. The solution is incubated for an additional 30 min at 37°C in the presence of 0.1% sodium dodecylsulphate and 50 µg proteinase K per ml. After repeated phenol extraction and ethanol precipitation, the DNA is dissolved in 10 mM Tris-HCl (pH 7.5).

Isolation and characterization of recombinant DNA clones containing trypanosome cDNA or nuclear DNA. The recombinant plasmids containing DNA complementary to VSG mRNAs inserted by GC tailing in the PstI site of plasmid pBR322 were isolated by Hoeijmakers et al. [9] and characterized previously [14,16,18]. MboI partial digestion products of trypanosome DNA were used to construct cosmid recombinant DNA libraries [26,27]. The cosmid vector used was a derivative of pJB8 and a gift from Dr. T.Lund. EcoRI sub-clones of the 118 BC cosmid were made in plasmid pAT153. Details of the procedure for constructing and screening the libraries will be described elsewhere.

Restriction endonuclease digestion, electrophoresis and transfer of DNA to nitrocellulose filters. Incubation conditions for restriction endonucleases were as specified previously [28].

DNA digests were size-fractionated by electrophoresis through horizontal agarose slab gels as described [29]. The completeness of nuclear DNA digestions was tested by adding bacteriophage lambda DNA to an aliquot of the incubation mixture, followed by size-fractionation in agarose gels and visual inspection for band stoicheiometry of the ethidium bromide-stained gel. After electrophoresis, the gel was incubated for 40 min in 0.25 M HCl to reduce the molecular weight of the DNA fragments [30]. The gel was then incubated in alkali, followed by neutralization and transfer of the DNA to nitrocellulose filters (Sartorius, pore size 0.1 μm) as described by Southern [31].

Isolation of DNA sub-probes and filter hybridization. Sub-probes of cloned DNA were isolated by preparative agarose gel electrophoresis of the appropriate restriction endonuclease digests, followed by solubilization of DNA-containing gel slices in saturated potassium iodide and adsorption of the DNA to activated glass as described [32]. The fragments were then labelled by nick-translation [33].

Hybridization of filters at 65°C in 3 x SSC (1 x SSC = 0.15 M NaCl, 0.015 M Na citrate (pH 7.0)), 10% dextransulphate was as described previously [30,34]. Fragments that might contaminate the probe were added in unlabelled form to the hybridization as competitor (probe:competitor = 1:2 weight ratio). Post-hybridizational washes to remove aspecifically-bound probe were for 2 h at 65°C with several changes of 3 x SSC. At this point the filters were either autoradiographed directly or washed further to melt off imperfectly matched hybrids by lowering the salt concentration stepwise (1 x, 0.3 x, 0.1 x SSC), each step taking 30 min at 65°C.

RESULTS

The physical maps of the BCs of the 117 and 118 VSG genes

Fig. 1 presents maps of the restriction endonuclease recognition sites in and around the 117 and 118 BC genes. The sites were located by standard techniques using cosmid recombinant DNA clones containing trypanosome nuclear DNA inserts of 30-40 kb. The co-linearity of the cloned DNA and the corresponding segment of nuclear DNA was verified by co-electrophoresis

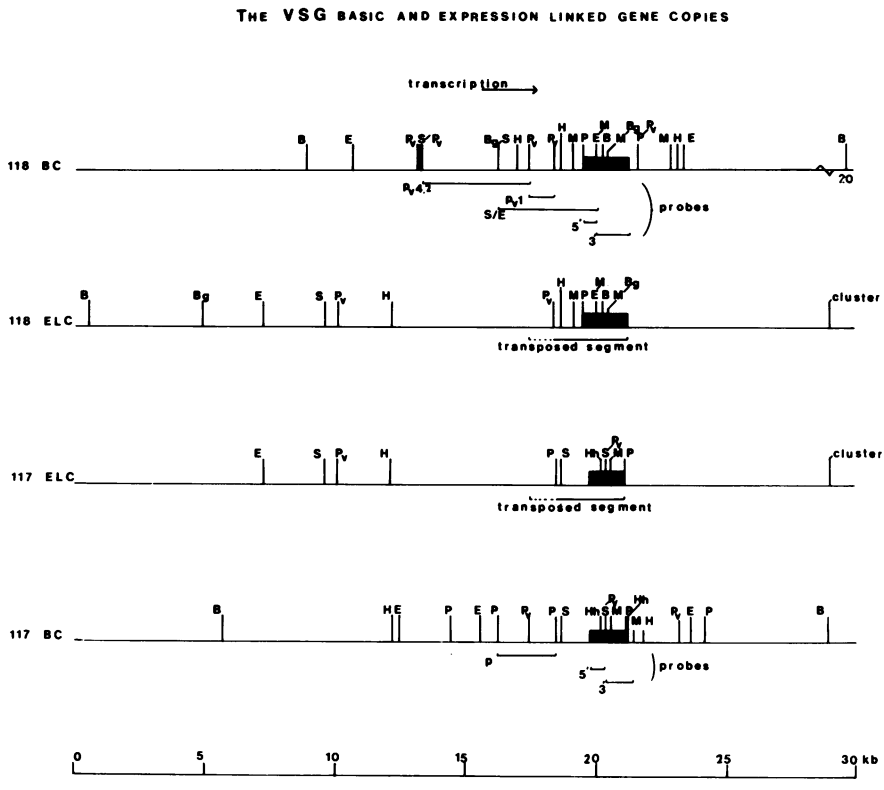

Fig. 1. Physical maps of the BC and ELC of the 117 and 118 VSG genes. The maps were constructed as described in the text. The black boxes correspond to the coding sequences. The DNA segments used as hybridization probes are indicated below the maps. Probes are designated by the restriction enzyme used to cut them from the BC gene clone and (if multiple fragments are used) by the fragment size in kb. The 5' and 3' half-gene probes were derived from the cDNA clones TcV-118.2 (cut with EcoRI) and TcV-117.5 (cut with SalI). The size of the transposed segment is indicated below the ELC map. The 'cluster' at the very 3'-end of the ELC maps indicates a cluster of restriction enzyme cleavage sites, including sites for HhaI, PvuI, HindIII, MspI, SalI, MboI and EcoRI. In the case of the 118 ELC, the cluster also contains recognition sites for BamHI, BglII and PstI. The 117 ELC cluster also contains sites for TaqI and HpaI. The MspI and HhaI physical maps of the 117 VSG gene are incomplete at the 5'-end of the gene. Abbreviations: B, BamHI; Bg, BglII, E, EcoRI; Hh, HhaI; H, HindIII, M, MspI, P, PstI; Pv, PvuII; S, SalI.

of restriction digests of cloned DNA and nuclear DNA (not shown). Proof that the cloned genes are indeed the BC genes and not related genes cross-hybridizing with the cDNA probes used has been presented elsewhere [18].

The physical map of the ELC of the 118 VSG gene

Since we have not succeeded in cloning an ELC gene, the physical maps of the ELCs had to be constructed by genomic blot analysis of total nuclear DNA. This analysis is complicated by the presence of related VSG genes that cross-hybridize with the cDNA probes used. The complication is relatively minor with the 118 gene, because the 118-related genes are few and sufficiently different from the 118 BC gene to allow the cross-hybridization to be removed by stringent washing. Fig. 2 shows blots used to construct the map of the 118 ELC and flanking regions. Each digest was done with nuclear DNA from variant 118 (lanes A) and variant 117 (lanes B). Hybridizing fragments present in both nuclear DNAs are derived from the 118 BC gene or 118-related genes. Hybridizing fragments only present in 118 nuclear DNA are derived from the 118 ELC gene. Endonuclease EcoRI cuts once within the 118 gene (see Fig. 1) and the left-hand panel of Fig. 2 shows that the 5' half-gene probe obtained by cutting the 118 cDNA with EcoRI recognizes a single 9-kb EcoRI fragment in heterologous nuclear DNA (lane B) and the same fragment plus

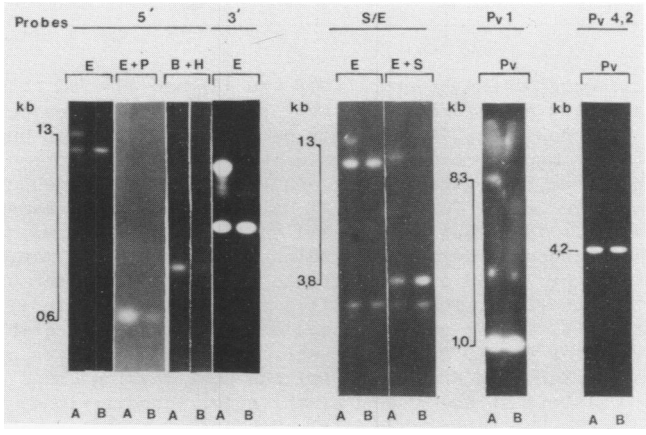

Fig. 2. Blot analysis of the 118 BC and ELC genes. Variants 118 (A) and 117 (B) nuclear DNA were digested with the restriction enzymes indicated, size-fractionated on agarose gels, blotted onto nitrocellulose filters and hybridized to ^{32}P-labelled DNA probes at 65°C. After post-hybridizational washes at 65°C down to 0.1 x SSC, the blots were exposed to X-ray film (Kodak XR-I). The TcV/118.2 probes and abbreviations are as in Fig. 1. See Methods for further details.

an additional 13 kb ELC fragment in homologous nuclear DNA (lane A). Both fragments are reduced to a single 0.6 kb fragment when the EcoRI digest is redigested with PstI. PstI, therefore, cuts in the transposed segment and the same holds for HindIII (Fig. 2). The middle panel of Fig. 2 shows blots hybridized with the SalI x EcoRI fragment (S/E) from the 118 BC gene clone (see Fig. 1 for the map position of this probe). In EcoRI digests of nuclear DNA this probe hybridizes with the same 9-kb and 13-kb fragments seen by the 5' half-gene cDNA probe. In a SalI-EcoRI double digestion the probe hybridizes to the shortened 10.2-kb ELC fragment and the 3.8-kb band from the 118 BC. In addition, it cross-hybridizes to a 3-kb fragment not derived from the 118 BC gene. Redigestion of the EcoRI digests with SalI reduced the size of the ELC fragment by 3 kb, placing the SalI site 10 kb upstream of the 118 ELC, as shown in Fig. 1.

In this way the cleavage sites around the 118 ELC were determined for a number of restriction enzymes. The resulting map, presented in Fig. 1, shows that the genomic environments of BC and ELC are completely different and that the genes are not detectably linked. Analysis of the cosmid clones containing the 118 BC gene has confirmed that this 118 BC is more than 35 kb away from the 118 ELC.

The size of the transposed segment

The approximate size of the 118 VSG gene segment that is transposed to create the 118 ELC follows from the comparison of the BC and ELC maps in Fig. 1. Upstream of the gene there are recognition sites for PvuII, HindIII and MspI which are present both in front of the BC and ELC genes, as inferred from the digestion of EcoRI- or BamHI-cut nuclear DNA with one of these enzymes (see Fig. 2).

The position of the 5' edge of the transposed segment was confirmed by additional hybridization experiments with the probes from the 118 BC gene clone indicated below the 118 BC gene map in Fig. 1. Fig. 2 shows that the Pv1 and S/E probes hybridize to extra fragments only present in 118 nuclear DNA, whereas the Pv-4.2 probe does not. The difference between BC and ELC maps nearest to the coding region is, therefore, the PvuII site 2 kb 5'

of the BC gene which is lacking in the ELC map. This locates the 5' edge of the transposed segment somewhere within the Pv1 fragment, which maps between 1.2 and 2.2. kb upstream of the gene.

The precise position of the 3' edge of the transposed segment was recently mapped within the 3'-terminal region of the 118 BC gene [18]. The transposed segment is, therefore, between 2.8 and 3.8 kb, i.e. about twice the size of the 1.7-kb 118 mRNA.

The physical map of the ELC of the 117 VSG gene

117 cDNA hybridizes to a large number of fragments in nuclear DNA and only part of these fragments disappear after stringent washing [11,15]. We attribute this to the presence of many 117-related VSG genes, some of which are sufficiently homologous to the 117 gene to cross-hybridize under stringent conditions with 117 cDNA probes. 3' half-gene probes recognize more fragments in blots than 5' half-gene probes, presumably because the 5' half of VSG genes evolves more rapidly than the 3' half [15]. The 117-related genes have considerably complicated the construction of the 117 ELC map drawn in Fig. 1.

Blots used for constructing the map are shown in Fig. 3. 5' half-gene probes mainly hybridize to fragments derived from the 117 BC and ELC, while fragments attributable to 117-related genes are only visible after longer exposure (panels I and II). BC and ELC gene fragments show up at approximately the same intensity in these blots. The 3' half-gene probes result in a more complex banding pattern, as illustrated in the right-hand part of panel I in Fig. 3. Since the enzymes used in this experiment, HindIII, EcoRI and BamHI, do not cut 117 cDNA, all bands recognized by the 5' probe also hybridize with the 3' probe. The relative intensity of the ELC fragments is often low after hybridization with the 3' probe. We attribute this to overlap of bands from multiple 117-related BC genes, distorting the ELC/BC stoicheiometry in the hybridization.

The 5' edge of the transposed segment is located in between the PstI site 1.3 kb and the PvuII site 2.4 kb upstream of the BC gene, because a 5' ELC fragment shows up after digestion of 117 nuclear DNA with PvuII and not with PstI. HindIII shortens the 5' PvuII fragment of the ELC by 2.1 kb and this locates

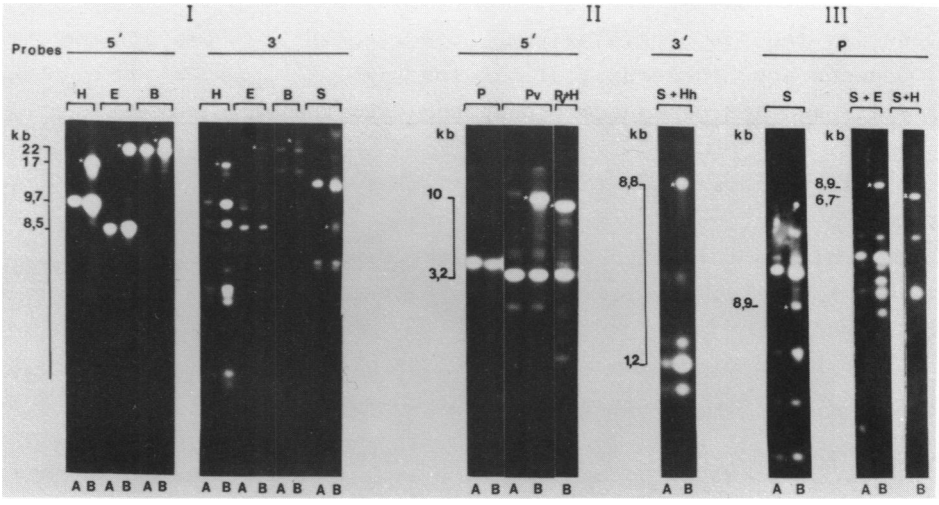

Fig. 3. Blot analysis of the 117 BC and ELC genes. The analysis was done as described in Fig. 2 with the 117 probes specified in Fig. 1. The ELC bands are indicated by asterisks. Abbreviations are as given in the legend to Fig. 1.

this HindIII site 5' of the 117 ELC (panel II of Fig. 3). The SalI site at this side of the ELC could not be located with the cDNA probes because there is a SalI site in the transposed segment outside the coding region (see Fig. 1). We therefore used the 117 BC clone to isolate the PstI fragment indicated in Fig. 1, which is located just upstream of the unaltered PstI site. This fragment contains the 5' edge of the transposed segment and it hybridizes to an 8.9-kb SalI fragment in 117 nuclear DNA that is absent in 118 nuclear DNA (Fig. 3, panel III). This fragment is not cut by EcoRI, whereas the large number of fragments derived from 117-related genes are greatly reduced in size by EcoRI (Fig. 3, panel III). In a SalI x HindIII double digest, on the other hand, no specific ELC fragment is visible because the fragments from 117 ELC and BC co-migrate.

The HhaI site just outside the coding region 3' of the BC gene (Fig. 1) is absent 3' of the ELC gene (see 8.8 kb fragment in the SalI x HhaI digest of Fig. 3). The transposed segment, therefore, must end close to the BC gene. The edge has recently

been located within the last 150 bp of the gene [18], as is the case for the 118 gene. The transposed segments of the 117 and 118 genes are, therefore, of similar size.

The 117 and 118 genes are transposed into the same expression site

Fig. 1 shows that the physical maps of the 117 and 118 ELCs are identical for all sites outside the transposed segment. This suggests that both genes use the same (or very similar) expression site. The following additional experiments support this conclusion:

1. After digestion of nuclear DNA with KpnI the 117 and 118 ELCs yield 21-kb fragments that co-migrate in gels (not shown).

2. Upstream of both ELCs there is a 6-kb DNA segment which is not cut by TaqI, as shown in the left-hand panel of Fig. 4. Probes that hybridize with the 5' edge of the transposed segment recognize analogous 6-kb ELC fragments in homologous and not in heterologous nuclear DNA. In the 118 ELC this segment is also not cut by Rsa I (not tested with 117). Since TaqI and RsaI cleave two different permutations of the tetranucleotide sequence AGCT, 6-kb DNA segments not cut by either of these enzymes must have a very unusual sequence.

Taken together these results show that the 117 and 118 genes are transposed into the same expression site (or different copies of the same site).

The ends of the transposed segment contain a repeated sequence

The 3' edge of the transposed segment lies within the last 150 bp of the BC gene [18]. This part of VSG genes is conserved in evolution [16,17] and DNA probes containing this sequence hybridize under non-stringent conditions to many bands in nuclear blots [12,13,15]. The middle and right-hand panels of Fig. 4 show that a repeated sequence is also present in the fragments from the 117 and 118 BC gene clones that overlap the 5' edge of the transposed segment. This hybridization is washed off in 0.1 x SSC at 65°C.

The 117 and 118 5' edge probes probably contain a similar repeated sequence because these probes cross-hybridize (not shown) and appear to recognize a similar set of fragments in

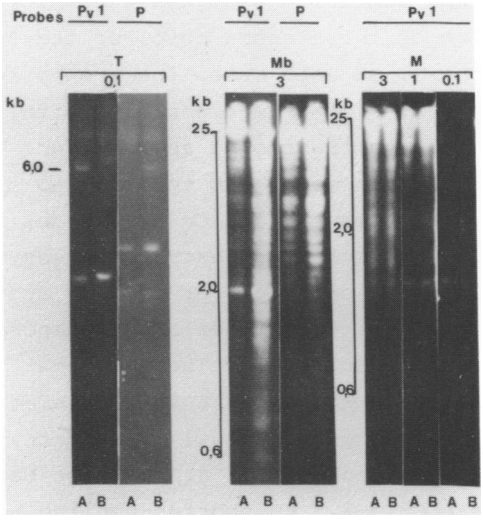

Fig. 4. Blot analysis of the 5' edge of the transposed VSG gene segments. Left part: TaqI (T) digests of 118 (A) and 117 (B) nuclear DNA hybridized under stringent conditions (see Fig. 2) with DNA segments containing the 5' edge of the transposed segments of the 118 and 117 VSG genes, i.e. the Pv1 fragment from the 118 BC clone and the P fragment from the 117 clone. Middle part: An analogous hybridization with MboI (Mb) digests of nuclear DNA, but washed only with 3 x SSC at 65°C. Right-hand part: MspI (M) digests of nuclear DNA after hybridization with the 118 PvuI 1.0-kb (Pv1) 5' breakpoint probe and post-hybridization washes with 3 x (3), 1 x (1) and 0.1 x (0.1) SSC at 65°C.

the MboI digest in Fig. 4. Unpublished experiments with cloned genomic DNA containing three other putative VSG genes (defined by their 3' homology with known VSG genes) show the repeated sequence to be present next to each of these genes. The 5' edge probes do not hybridize to DNA fragments containing the 3'-ends of VSG genes, however. Different repeated sequences are, therefore, present at each end of the transposed sequence.

DISCUSSION

The early maps of the 118 VSG genes reported by Borst et al. [12,15] established three basic characteristics of the expression of VSG genes: Expression is accompanied by a gene duplication, the duplicated gene is transposed into new surroundings and the transposed segment includes at least 1 kb in front of the gene. The more detailed maps for the 117 and 118

genes presented here confirm and extend this picture and show that the expression of these genes involves the same duplication-transposition process. In both cases the transposed segment starts 1-2 kb in front of the gene and ends within the last 150 bp of the gene [18]. The transposed segments are inserted into a 14-kb region of the genome, which contains no sites for any of the restriction enzymes tested. On both sides this 14-kb region is flanked by a series of restriction enzyme cleavage sites that appear to be identical for the 117 and 118 ELC. We conclude, therefore, that both genes are transposed into the same site (or very similar sites). The fact that the position of both genes in the expression site is the same, is probably accidental. In a second, independently-cloned trypanosome expressing the 118 VSG gene (MITat 1.176), the barren region downstream of the ELC is 2.7 kb shorter and the upstream barren region is 10 kb longer (Michels, P.A.M. and Van der Ploeg, L.H.T., unpublished). The expression site may, therefore, undergo sequence rearrangements or there may be several slightly different sites.

An analogous study was recently carried out by Pays et al. [24a]. From mapping experiments with nuclear DNA blots they infer that the transposed segments in the Antat 1.1 and 1.8 genes contain 0.2-0.8 kb upstream of the gene. Two independent clones expressing the 1.1 VSG gene were found to give analogous but non-identical ELCs. The authors 'were unable to obtain a clear picture of the 3' flanking regions' [24a] and no site in this region was precisely mapped. It is, therefore, not possible to judge whether the expression site in the Antat stock is similar to the one illustrated in Fig. 1.

We have no information on the nature of the unusual DNA surrounding the transposed genes in the expression site. We do not know whether the sequences on both sides are similar or whether they occur also elsewhere in the genome. The absence of restriction enzyme cleavage sites suggests that these sequences are AT-rich, repetitive or both, but they are not made up of complete repeats of the AT-rich 177-bp satellite DNA of *T. brucei*, because this is cleaved by endonuclease HhaI [35]. Whereas the 'barren' sequence is bordered at the 5'-side by a

fairly conventional looking piece of DNA, the 3' border is formed by a remarkable cluster of restriction sites. Within the limits of error of our experiments, these sites all map in the same position. Rather than a site cluster, this may be the end of a chromosome.

Our experiments also provide information on possible mechanisms for the replicative transposition that gives rise to the ELC. We have previously proposed that the 3'-end of VSG genes recombines with homologous sequences in the expression site and that the 3'-end of the gene is replaced in this recombination by sequences contributed by the expression site [18]. This proposal explains the conserved nature of the 3' untranslated sequence of VSG mRNAs [16,17] and the absence from the 117 and 118 BC genes of the 3'-terminal sequence of the corresponding mRNA [18].

In this paper we have shown that fragments overlapping the 5'-end of the transposed segment also contain a sequence that is present in many copies in the genome. This conserved sequence (which differs from the one present at the 3'-end of VSG genes) could mediate a recombination between the 5'-end of the transposed sequence and the expression site. The duplicative transposition of VSG genes may, therefore, involve a unidirectional gene conversion in which the initial pairing of the BC gene and expression site depend on limited homology between the ends of the transposed sequence and sequences present in the expression site, e.g. the preceding ELC.

The mechanism envisaged here for the switching of surface antigens in trypanosomes is similar to the mechanism of the mating-type switch in yeast [36,37]. Mating-type is controlled by the MAT locus, which may contain a or α information. Transposition of a copy of either of two silent genes - called HML and HMR - may change the gene at MAT from a to α or α to a. HML, HMR and MAT are flanked on both sides by long identical sequences that are thought to mediate a unidirectional gene conversion, switching the information at MAT [38-41]. Transcription starts in the middle of the transposed segment and the suppression of transcription of a or α sequences at the silent loci HML and HMR is thought to be the result of local

chromatin structure induced by the products of the SIR (= MAR) genes, which act on a DNA segment outside the transcribed region [39,41].

Although the similarity between the control of antigenic variation and mating-type is obvious, it should be noted that the differences between these systems are not limited to the number of silent genes, which is 2 in yeast and more than 100 in trypanosomes. Whereas the homology between the silent cassettes and expression site includes several hundred base pairs on both sides of the transposed sequence in yeast [38, 40], homology in trypanosomes is limited to several sequence blocks of 10-15 bp, at least at the 3'-end of the transposed sequence [17]. Whereas mating-type switches after each cell division with more than 70% chance to the opposite type [42], antigenic switches in trypanosomes are much less frequent [2,3]. Finally, our preliminary data on VSG gene transcription suggest that the control of VSG gene expression may be radically different from that of mating-type genes. We have found minor variant-specific transcripts hybridizing with the 1-kb segment in front of the VSG gene and, moreover, S_1 nuclease protection experiments (and sequence comparison of 118 BC and 118 cDNA clones) show that the 118 BC gene lacks the 5'-end of the mature mRNA [18; Van der Ploeg, L.H.T., Liu, A.Y.C. and Bernards, A., unpublished]. In the speculative model that accounts for these results transcription starts either outside the transposed segment or near its border. The primary transcript is spliced to yield a mature mRNA joining sequences that are discontinuous in the DNA. If this interpretation is correct, transcription of BC VSG genes may not be actively suppressed, as are the silent copies of a and α information in yeast. Rather, transcription may be activated by the transposition, either by bringing the gene under a promoter that pre-existed in the expression site or by creating a new promoter at the border of the transposed segment.

Why do trypanosomes use this complex duplication-transposition mechanism to control the expression of their VSG gene repertoire? As discussed more fully elsewhere [43], it seems likely that this system has evolved to ensure that only one VSG

gene is expressed at a time in an imprecisely programmed order [3,44]. In its simplest form such a system has a single expression site (or two sites with a rigorous system of allelic exclusion) and the VSG genes can only be expressed after transposition into this site. Whether there is only a single expression site remains to be seen, but the prediction that VSG genes can only be expressed via duplication-transposition is an oversimplification. There is a class of genes expressed early in infection to which our 221 gene and three of the four VSG genes studied by Williams and co-workers belong [8,19-21], that do not seem to be either duplicated or transposed when activated. How the expression of these genes is coördinated with the expression of genes that give rise to an ELC, remains to be settled.

ACKNOWLEDGEMENTS

We thank Dr. A.C.C.Frasch (Catedra de Bioquímica, Facultad de Medicina, Universidad de Buenos Aires, Argentina) for help in some of these experiments and Mrs J.Van den Burg for technical assistance. We are grateful to all collaborators of the Laboratory of Gene Structure and Expression (National Institute for Medical Research, Mill Hill, London NW7 1AA, UK), especially to Drs F.G.Grosveld and R.A.Flavell for advice on cosmid isolation and to Dr. T.Lund for the gift of the pJB8 derivative cosmid cloning vector. We thank Drs J.C.Boothroyd and G.A.M.Cross (The Wellcome Research Laboratories, Department of Immunochemistry, Langley Court, Beckenham, Kent BR3 3BS, UK) and Dr. R.Benne for helpful comments on the manuscript. This work was supported in part by a grant to PB from the Foundation for Fundamental Biological Research (BION) which is subsidized by The Netherlands Organization for the Advancement of Pure Research (ZWO).

Abbreviations: kb, kilo-base pair(s); VSG, variant surface glycoprotein; mRNA, messenger RNA; cDNA, complementary DNA; ELC, expression-linked copy; BC, basic copy; NET, 100 mM NaCl, 100 mM EDTA, 10 mM Tris-HCl (pH 7.5); SSC, 0.15 M NaCl, 0.015 M Na citrate (pH 7.0).

REFERENCES

1 Cross, G.A.M. (1975) Parasitol. 71, 393-417.
2 Cross, G.A.M. (1978) Proc.Roy.Soc.London Ser. B 202, 55-72.
3 Vickerman, K. (1978) Nature 273, 613-617.
4 Eggitt, M.J., Tappenden, L. and Brown, K.N. (1977) Parasitol. 75, 133-141.
5 Williams, R.O., Marcu, K.B., Young, J.R., Rovis, L. and Williams, S.C. (1978) Nucl.Acids Res. 5, 3171-3182.
6 Lheureux, M., Lheureux, M., Vervoort, T., Van Meirvenne, N. and Steinert, M. (1979) Nucl.Acids Res. 7, 595-609.

7 Cordingly, J.S. and Turner, M.J. (1980) Mol.Biochem.Parasitol. 1, 129-137.

8 Williams, R.O., Young, J.R. and Majiwa, P.A.O. (1979) Nature 282, 847-849.

9 Hoeijmakers, J.H.J., Borst, P., Van den Burg, J., Weissmann, C. and Cross, G.A.M. (1980) Gene 8, 391-417.

10 Pays, E., Delronche, M., Lheureux, M., Vervoort, T., Bloch, J., Gannon, F. and Steinert, M. (1980) Nucl.Acids Res. 8, 5965-5981.

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

12 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., Suppl. 29, 1033-1036.

13 Frasch, A.C.C., Bernards, A., Van der Ploeg, L.H.T., Borst, P., Hoeijmakers, J.H.J., Van den Burg, J. and Cross, G.A.M. (1980) In: The Biochemistry of Parasites and Host-Parasite Relationships: The Host-Invader Interplay (Van den Bossche, H., Ed.), North-Holland, Amsterdam, pp. 235-239.

14 Boothroyd, J.C., Cross, G.A.M., Hoeijmakers, J.H.J. and Borst, P. (1980) Nature 288, 624-626.

15 Borst, P., Frasch, A.C.C., Bernards, A., Van der Ploeg, L.H.T., Hoeijmakers, J.H.J., Arnberg, A.C. and Cross, G.A.M. (1981) Cold Spring Harbor Symp.Quant.Biol. 45, 935-943.

16 Boothroyd, J.C., Paynter, C.A., Cross, G.A.M., Bernards, A. and Borst, P. (1981) Nucl.Acids Res. 9, 4735-4743.

17 Majumder, H.K., Boothroyd, J.C. and Weber, H. (1981) Nucl.Acids Res., 9, 4745-4753.

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

19 Williams, R.O., Young, J.R., Majiwa, P.A.O., Doyle, J.J. and Shapiro, S.Z. (1980) Am.J.Trop.Med.Hyg., Suppl. 29, 1037-1042.

20 Williams, R.O., Young, J.R., Majiwa, P.A.O., Doyle, J.J. and Shapiro, S.Z. (1981) Cold Spring Harbor Symp,Quant.Biol. 45, 945-949.

21 Marcu, K.B. and Williams, R.O. (1981) In: Genetic Engineering (Setlow, J.K. and Hollaender, A., Eds), Vol. 3, Plenum, New York, pp. 129-155.

22 Agabian, N., Thomashow, L., Milhausen, M. and Stuart, K. (1980) Am.J. Trop.Med.Hyg., Suppl. 29, 1043-1049.

23 Pays, E., Van Meirvenne, N., LeRay, D. and Steinert, M. (1981) Proc. Natl.Acad.Sci.U.S. 78, 2673-2677.

24 Pays, E., Lheureux, M. and Steinert, M. (1981) Nature 292, 265-267.

24a Pays, E., Lheureux, M. and Steinert, M. (1981) Nucl.Acids Res. 9, 4225-4238.

25 Fairlamb, A.H., Weislogel, P.O., Hoeijmakers, J.H.J. and Borst, P. (1978) J.Cell Biol. 76, 293-309.

26 Grosveld, F.G., Dahl, H.H.M., De Boer, E. and Flavell, R.A. (1981) Gene 13, 227-237.

27 Collins, J. and Hohn, B. (1978) Proc.Natl.Acad.Sci.U.S. 75, 4242-4246.

28 Borst, P. and Fase-Fowler, F. (1979) Biochim.Biophys.Acta 565, 1-12.

29 Borst, P., Fase-Fowler, F. and Gibson, W.C. (1981) Mol.Biochem.Parasitol. 3, 117-131.

30 Wahl, G.M., Stern, M. and Stark, G.R. (1979) Proc.Natl.Acad.Sci.U.S. 76, 3683-3687.

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

32 Vogelstein, B. and Gillespie, D. (1979) Proc.Natl.Acad.Sci.U.S. 76, 615-619.

33 Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J.Mol.Biol. 113, 237-251.
34 Jeffreys, A.J. and Flavell, R.A. (1977) Cell 12, 429-439.
35 Borst, P., Fase-Fowler, F., Frasch, A.C.C., Hoeijmakers, J.H.J. and Weijers, P.J. (1980) Mol.Biochem.Parasitol. 1, 221-246.
36 Oshima, Y. and Takano, I. (1971) Genetics 67, 327-335.
37 Hicks, J.B., Strathern, J.N. and Herskowitz, I. (1977) In: DNA Insertion Elements, Plasmids, and Episomes (Bukhari, A., Shapiro, J. and Adhya, S., Eds), Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., pp. 457-462.
38 Nasmyth, K.A. and Tatchell, K. (1980) Cell 19, 753-764.
39 Nasmyth, K.A., Tatchell, K., Hall, B.D., Astell, C. and Smith, M. (1981) Nature 289, 244-250.
40 Strathern, J.N., Spatola, E., McGill, C. and Hicks, J.B. (1980) Proc.Natl.Acad.Sci.U.S. 77, 2839-2843.
41 Klar, A.J.S., Strathern, J.N., Broach, J.R. and Hicks, J.B. (1981) Nature 289, 239-244.
42 Strathern, J.N. and Herskowitz, I. (1979) Cell 17, 371-381.
43 Borst, P. and Cross, G.A.M. (1982) Cell, in press.
44 Capbern, A., Giroud, C., Baltz, T. and Mattern, P. (1977) Exptl. Parasitol. 42, 6-13.