Journal of Cellular Biochemistry 24:287–295 (1984)
Molecular Biology of Host–Parasite Interactions 145–153

The Role of Duplication in the Expression of a Variable Surface Glycoprotein Gene of Trypanosoma brucei

John R. Young, Mervyn J. Turner, and Richard O. Williams

MRC Biochemical Parasitology Unit, The Molteno Institute, Cambridge CB2 3EE, UK (J.R.Y., M.J.T.) and International Laboratory for Research on Animal Diseases, P.O. Box 30709, Nairobi, Kenya (R.O.W.)

The variable surface glycoprotein (VSG) genes of Trypanosoma brucei have been classified into two groups depending upon whether or not duplication of the genes is observed when they are expressed. We report here the observation of duplication apparently linked to espression of the ILTaT 1.3 gene in the ETaR 1 trypanosome stock. In the ILTaR 1 stock, expression of the ILTaT 1.3 VSG did not involve a new duplication, but instead activation of a preexisting gene copy that had been apparently generated earlier by a duplication event analogous to that directly observed in the ETaR 1 trypanosomes. The results suggest that the well-characterised gene duplications found with other VSG genes are common to all VSG genes but are not directly responsible for controlling expression. All currently available data can be accomodated by a model that assumes that gene duplication and replacement occurs independently of antigenic switching.

Key words: Trypanosoma brucei, variable surface glycoprotein, gene duplication

Data apparently showing strict linkage between the presence of a duplicated copy of some VSG genes and their expression [1] have led to the implication that this gene duplication itself controls the expression of VSGs [2]. Since other VSG genes are not duplicated when expressed [3–6], this implies that there are two classes of VSG genes controlled by different mechanisms [6]. A detailed study of the ILTaT 1.3 VSG gene showed that expression was not linked to duplication, but that the site occupied by the expressed copy was similar to those occupied by expression-linked copies (ELCs) [5]. Analysis of the gene in other trypanosome stocks suggested that the expressed copy (called copy C) of the ILTaT 1.3 gene had arisen by two successive duplication events, although these were not directly observed. The first event was similar to that observed in the production of ELCs of other VSG genes. A similar region of a nontelomeric copy (copy A), analogous to the basic copy (BC) which is duplicated to produce an ELC, was duplicated and inserted into a telomeric site. The expressed copy, however, appeared to have arisen by a second duplication of the first

Received March 10, 1983; revised and accepted December 2, 1983.

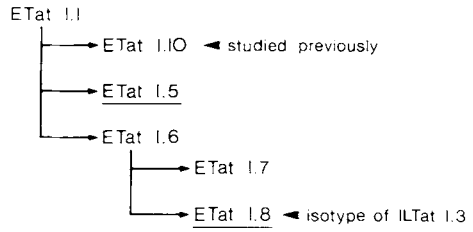

Fig. 1. Relationship of ETaR 1 clones [9]. Each clone was isolated from a relapse population of an infection with the preceding clone. The six clones shown were used in the work reported here. Clones ETaT 1.5 and ETaT 1.8 (underlined) were used in the experiments shown. ETaT 1.8 expressed a VSG serologically indistinguishable from ILTaT 1.3.

telomeric copy (copy D), which, unlike the preceding event, involved duplication of sequences beyond the 3′ end of the gene [5]. Potentially analogous telomere-to-telomere duplication has recently been observed for the expression-linked duplication of other VSG genes [7,8]. A clone from the ETaR 1 trypanosome stock contained copies of the ILTaT 1.3 gene that were identical in internal and flanking restriction enzyme sites to copies A and D in the ILTaT trypanosomes. The expressible copy (copy C), which was present in all ILTaR 1 clones whether or not ILTaT 1.3 was expressed, was absent in the ETaR 1 clone.

We have investigated the genomic changes that occur in the clone ETaT 1.8, which expresses a VSG serologically indistinguishable from ILTaT 1.3. In this clone we have found a newly duplicated copy of the ILTaT 1.3 gene, which has arisen by duplication of the other telomeric copy (copy D). This result supports our previous inference that the expressed copy in the ILTaT 1.3 clone was derived by duplication of the other telomeric copy. These observations suggest that whether or not expression linked duplication is observed it is not a characteristic of particular VSG genes. The differing observations that have been obtained with different VSG genes can be accommodated within a single model for activation and expression of all VSG genes.

MATERIALS AND METHODS

ETaR 1 trypanosome clones were provided by Dr A.R. Gray. Their pedigree is described in detail elsewhere [9]. The sequential relationship of the clones used in this work are shown in Figure 1. Antigenic homegeneity of the ETaT 1.8 population was confirmed by immunofluorescence with anti-ETaT 1.8 and anti ILTaT 1.3 antisera.

The ILTaR 1 trypanosome clones have been described elsewhere [5]. Trypanosome clones are referred to by the standard nomenclature for the VSGs they express. Cloned stocks are referred to by the repertoire of VSGs they are capable of expressing. Thus ETaR 1 and ILTaR 1 are, respectively, Edinburgh Trypanozoon antigen Repertoire No. 1 and ILRAD Trypanozoon antigen Repertoire No. 1, and, for example, ETaT 1.5 is Edinburgh Trypanozoon antigen Type 5 from repertoire No. 1. ETaR 1 and ILTaR 1 stocks have extensively overlapping antigen repertoires (S.Z. Shapiro, personal communication) and have nearly identical genomes as judged by Southern blot hybridisation with VSG cDNA and other probes (Massamba et al, in press). They were isolated four years apart (1960, 1964 respectively) from the same area in East Africa [9,10].

Construction of the plasmid pcBCl, containing the entire coding sequence for ILTaT 1.3 VSG, has been described [3]. The cDNA sequence has been determined

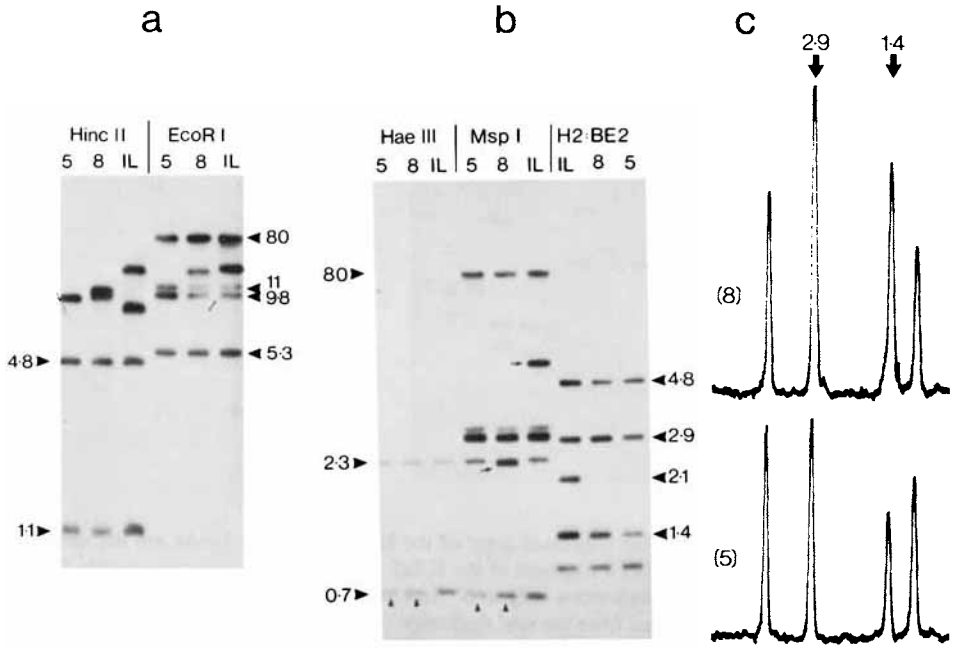

Fig. 2. Expression-linked duplication of the ILTaT 1.3 gene. 1a) Hybridisation of the whole ILTaT 1.3 cDNA to Hinc II and EcoR I digests of DNA from trypanosome clones ETaT 1.5 (5), ETaT 1.8 (8), and ILTaT 1.4 (IL). Neither of these enzymes cuts the cDNA sequence. b) Hybridisation of a probe lacking the 3′ 450 bp of the cDNA to the same three DNAs digested with Hae III, Msp I, or with both Hinc II and BstE II (H2:BE2). Internal 700-bp Msp I and Hae III fragments are marked with arrowheads. Small arrows indicate 5′ Msp I fragments present in ETaT 1.8 and ILTaT 1.4 that are absent from ETaT 1.5. The source of various fragments in indicated in the map of Figure 4. Sizes of fragments are marked in kilobase pairs. c) Densitonemetric scans of the right-hand two tracks in b.

[11]. Methods for isolation of trypanosome DNA, preparation of cDNA probes and Southern blot hybridisation were as previously described [3]. A rough quantitation of hybridisation in the blot shown in Figure 2b was obtained using a Joyce-Loebl scanning densitometer (Mark III CS).

RESULTS

DNA isolated from the ETaR 1 clones shown in Figure 1 was used in Southern blot hybridisations with ILTaT 1.3 cDNA probes. Results obtained with all clones except ETaT 1.8 were similar to those described for ETaT 1.10 [5], and may be summarised in the statement that these clones contained four of the five copies of the gene present in all ILTaR 1 clones [5], including the proposed intermediate telomeric copy (D), but not the expressed copy (C). No differences were found between these ETaR 1 clones, with the exception of the distance between the copy D gene and the adjacent chromosome end. This variation in the 3′ flanking telomeric fragments is a well-documented feature of telomeric VSG genes [3,12]. All restriction enzyme sites mapped in and around all copies present in the ETaT trypanosomes were identical to those of the corresponding gene copies in the ILTaR 1 trypanosomes. The same enzymes failed to cut within a measurable distance of the 5′ end of copy D in both the ILTaT and ETaT clones. Thus the identity of copies A,B1, B2 and D [5] of the

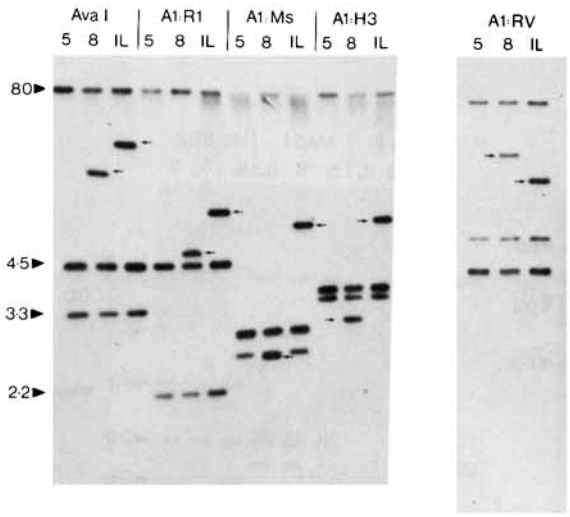

Fig. 3. 5′ flanking fragments of the expressed copy of the ILTaT 1.3 gene. DNAs are the same as Figure 2. The probe used was an Ava I fragment of the ILTaT 1.3 cDNA clone containing only the 5′ 450 bp of the cDNA sequence. It hybridises only to 5′ flanking fragments in the digestions used [5]. Arrows indicate the fragment derived from the new duplicated copy in ETaT 1.8 DNA (8) and from the expressed copy in the ILTaT DNA (IL) [5]. A1 = Ava I, R1 = EcoR I, Ms = Msp I, H3 = Hind III, RV = EcoR V. Al: R1, etc, indicates a double digestion with the two enzymes.

gene present in the ILTaT and ETaT clones, and of their flanking sequences, is established as far as is possible without complete sequence data. ETaT 1.5 DNA is used as representative of these clones in the experiments shown here. With the enzymes Hinc II and EcoR 1, which do not cut the cDNA sequence, there are four fragments that hybridise to the whole cDNA probe in the nonexpresser ETaT 1.5 DNA (Fig. 2a). The 1.1-kilobase (kb) Hinc II fragment is a doublet of identical fragments [5]. However, in the clone ETaT 1.8, which expressed a VSG seroligically identical to ILTaT 1.3, an additional fragment is seen with either enzyme. This fragment is comparable in size to the range observed for the expressed copy fragment in ILTaR 1 clones [5] (Fig. 2a).

The source of the duplicated copy in ETaT 1.8 was investigated in the experiments shown in Figures 2b,c. In the Hinc II BstE II double digests, no different bands appear in the ETaT 1.8 DNA compared to the ETaT 1.5 DNA. However, the 2.9-and 1.4-kb bands are approximately double in relative intensity, as shown in the scan (Fig. 2c). These two fragments extend from the single internal BstE II site to point 1.7 kb 3′ and 1 kb 5′ of the telomeric copy gene (D) respectively [5] (see map in Fig. 4). Internal 700-base pair (bp) Hae III and Msp I fragments which extend near to the 3′ end of the cDNA (Fig. 4) are also conserved in the duplication (Fig. 2b), as is the Ava II site close outside the 3′ end of the gene (not shown). These fragments are not present in copies A, B1, or B2 [5]. We conclude that the newly duplicated copy of the gene in the ETaT 1.8 genome arose by duplication of the telomeric copy D.

A new fragment was observed in the Msp I digest of ETaT 1.8 DNA, which was presumed to be a 5′ flanking fragment of the newly duplicated copy (Fig. 2b, small arrow), although it differed in size from the 5′ flanking Msp I fragment of the

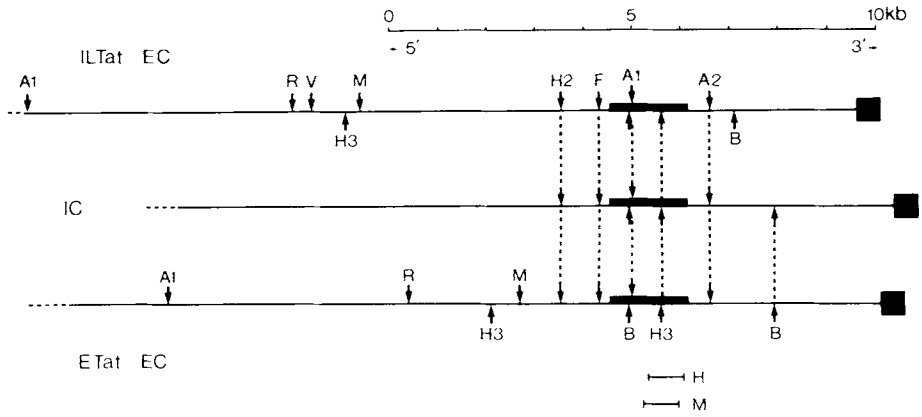

Fig. 4. Restriction maps of the telomeric copies of the ILTaT 1.3 gene in ETaR 1 and ILTaR 1 clones. The intermediate copy (IC) is present and identical in all ETaR 1 and ILTaR 1 clones studied. The ILTaT expressed copy (ILTaT EC) is present in all ILTaR 1 clones (5), but hte ETaT EC was found only in ETaT 1.8. The thickened region indicates the position of the cDNA sequence. Large blocks at the right-hand end are the adjacent chromosome ends (distance not to scale). Dotted lines connect enzyme sites common to different copies. The two bars below the maps show the positions of the internal 700 bp Hae III (H) and Msp I (M) fragments common to all three gene copies. Al = Ava I, A2 = Ava II, B = BstE II, F = Hinf I, H2 = Hinc II, H3 = Hind III, M = Msp I, R = EcoR 1, V = EcoR V.

expressed copy in the ILTaR 1 clones (Fig. 2b, small arrow). Sites beyond the conserved Hinc II site 5′ of the duplicated copy were mapped using a 450-bp 5′ Ava I fragment of the cDNA to probe double digests with Ava I and EcoR I, Msp I, Hind III, and EcoR V. The new fragments are easily recognised by comparison with ETaT 1.5 DNA (Fig. 3). Of the enzymes used only EcoR V did not cut between the gene and the 5′ flanking Ava I site.

The data from these experiments is summarised in the map shown in Figure 4. The intermediate copy D (labelled IC), is indistinguishable in the ETaR 1 and ILTaR 1 clones, and is present in all of them. ILTaT EC is the copy, which is expressed in ILTaR 1 clones expressing ILTaT 1.3 VSG. ETaT EC is the newly duplicated copy found only in ETaT 1.8 among the ETaR 1 clones, and whose presence appears therefore to be linked to the expressions of the VSG isotypic with ILTaT 1.3. The large blocks represent the putative chromosome ends. Sites common to the different gene copies are connected by dotted lines.

DISCUSSION

We have shown that expression of a VSG indistinguishable from ILTaT 1.3 in the ETaR 1 serodeme is accompanied by the appearance of a newly duplicated copy of the ILTaT 1.3 gene. This duplication differs from most previously described expression-linked duplications in two ways. The substrate of the duplication is telomeric, and the 3′ limit of the duplicated segment lies outside the 3′ end of the mRNA sequence. Since no further sites were found between the conserved BstE II site 1.7-kb 3′ of the gene and the adjacent chromosome end (Fig. 4), it is possible that the entire region up to the end of the chromosome was duplicated and transposed. Comparable expression-linked duplication of telomeric VSG genes has been observed

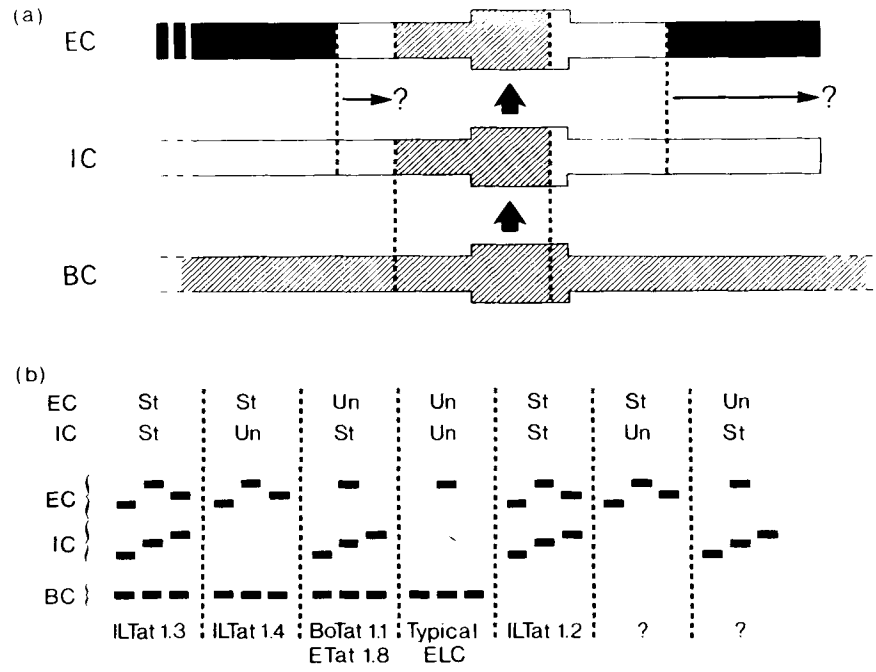

Fig. 5. a) Model for two-stage duplication giving rise to expressed copies of VSG genes. BC is a nontelomeric basic copy gene. IC is the intermediate telomeric copy and EC the telomeric expressed copy. The large vertical arrows represent duplication of the segment indicated by the surrounding dotted lines, and replacement of the preexisting sequences at the accepting site. Horizontal arrows and question marks represent possible variation in the location of the ends of the second duplication. The locations may be different in different events. The thicker region contains sequences found in the VSG mRNA. Shading represents the source of the sequences that are found at the expression site: shaded, basic copy; white, intermediate site; black, expression site. The model proposes the coexistence of several intermediate and expression sites undergoing these processes independently. b) Different patterns of hybridisations predicted by the model. Each set of three tracks represents an experiment in which a clone preceding expression, a clone expressing, and a clone following relapse are probed with the VSG gene concerned. EC and IC are shown as variable fragments typical of telomeric copies. Their presence depends upon the frequency or chance of replacement at the corresponding sites. Frequent replacement makes the gene copy appear unstable (Un). Infrequent replacement makes it stable (St). The three right-hand patterns are expected when the basic copy gene has been lost (see text). Examples of genes where these patterns have been observed are indicated below each. References are ILTaT 1.3 [5], ILTaT 1.4 [4], BoTaT 1.1 [8], Typical ELC [6], ILTaT 1.2 [14].

recently by others [7,8], but it was not determined whether recombination within the 3′ end of the gene, as observed for other BC to ELC duplications [6], occurred.

This duplication was analogous to that which, while not directly observed, had apparently occurred in the ILTaR 1 stock before the isolation of the originating clone ILTaT 1.1. This supports our previous conclusion that the copy expressed in the ILTaT 1.3 clones, copy C, arose by duplication of the telomeric intermediate copy (D) in these trypanosomes also. On the 5′ side of the duplicated segment in both trypanosome stocks at least 1 kb 5′ of the gene is cotransposed with the duplicated gene (Hinc II site, Fig. 4). This is similar to the 5′ limit of duplciated segments common to basic (BC) and expression-linked (ELC) copies of other VSG genes [6].

Figure 5a is a diagrammatic representation of the two successive duplications by which the expressed copies of the ILTaT 1.3 gene appears to have arisen in the two trypanosome stocks. (BC = copy A, basic copy [5]; IC = intermediate copy D [5]; EC = expressed copy, copy C). The first event, which was implied but not directly observed [5], is very similar to that reported for the generation of ELCs from BCs. The product, however, is not the copy eventually expressed, but is the precursor for the second duplication, which has been observed directly in the work reported here. In the second event the duplication extends from a similar, but not necessarily identical point 5′ of the gene to a point beyond the 3′ end possibly including the chromosome end. The product is an expressible copy, but its presence is not sufficient to cause expression.

In the trypanosomes we have studied, the first telomeric copy, copy D, plays the role of an intermediate copy (IC) between the nontelomeric basic copy (BC), copy A [5], and the expressible copy (EC). It is telomeric and located within a region barren of restriction enzyme sites, as in all expressed VSG genes, but its location on a minichromosome and the exceptional length of the 5′ barren region [12], distinguish it from other expressed copies. It is possible that it had been expressed at some time before the isolation of these trypanosome stocks, but there is no evidence as to whether this is so, or whether the role of copy D is solely as an intermediate.

VSG genes have been divided into two groups depending upon whether or not expression-linked duplication is observed [6]. Observations with some genes (the ELC genes) are consistent with the direct involvement of duplication in the control of expression [1], whereas observations with the other group of VSG genes are not [5,6]. In the ILTaR 1 trypanosomes the ILTaT 1.3 gene was expressed by transcriptional activation of a preexisting, ELC-like, telomeric copy [5] and therefore fell into the latter group. In the ETaR 1 trypanosomes, however, where the corresponding expressible gene copy was absent, expression was accompanied by the appearance of a newly duplicated telomeric gene copy, placing the gene in the ELC group. Thus the two kinds of observation characteristic of the two gene classes are not exclusively associated with particular VSG genes, and the division of VSG genes into the two classes is questionable. Another explanation for the differences in observations is then required.

Since several expressible copies of different VSG genes may coexist in the genome at different telomeric sites, while only one of them is expressed [5,15], there must exist a mechanism, as yet unidentified, ensuring mutually exclusive expression at these sites. Those sites at which ELCs have been observed, and those at which other VSG genes have been seen to be activated in the absence of duplications may be essentially equivalent and subject to the same mutually exclusive control mechanism. Duplicative replacement of VSG genes in these multiple expression sites may then be considered as a mechanism for the promotion of genes from other sites into the pool of expressible copies [3] and not the mechanism of selective VSG gene expression. Antigenic switching may occur either by replacement of the VSG gene at the currently active site, or by a switch to expression from a different site. In the former case, a newly duplicated "ELC" will appear, and in the latter a preexisting telomeric gene will be expressed.

There is not a priori reason to suppose that gene replacement occurs only during antigenic switching. Different genes and different sites may be subject to gene replacement with different frequencies. Trypanosome populations compared either

side of an antigenic switch, are separated by two highly selective events, the relapse itself and cloning, each of which is followed by expansion through many cell generations. Thus the effect of more or less frequent gene replacements will be manifested as an apparent instability or stability of particular telomeric gene when such populations are compared. This differential stability of particular telomeric gene copies may account for the preponderant association of one or other of the alternative observations with particular VSG genes.

If the presence or absence of particular gene copies is the result of susceptibility to gene replacement as well as the chance of selecting of particular clones, then the various observations that have been obtained with different VSG genes can be accounted for by the scheme of gene activation by two successive duplications, as in Figure 5a. Figure 5b shows the different observations that are predicted when the absence of intermediate and expressable gene copies in cloned populations is a result of susceptibility of their sites to replacements or the chance of clonal selection. The two patterns at the right would be observed when the intermediate copy is stable, but the original, nontelomeric, basic copy has been lost. Occasional loss of basic copies is to be expected if basic copies themselves evolve by gene duplication and subsequent divergence [5,14]. Thus the different kinds of observation can be accounted for within a single scheme of activation and selection.

This model suggests the possible involvement of an intermediate telomeric copy in the generation of all expressed copies. The second duplication as shown in Figure 5a carries 3′ flanking sequences from the intermediate site, and may also carry 5′ flanking sequences, into the expression site. If such an intermediate duplication is involved in the generation of ELCs, then the new 3′ end attached to the VSG gene segment derived from the basic copy must be derived, by duplication, from the intermediate telomeric site. If no intermediate is involved, then the new 3′ end will consist of sequences already present in the expression site. This difference provides a possible test for the hypothetical involvement of an unstable intermediate telomeric copy. Similarly at the 5′ end, newly duplicated sequences should sometimes be found that are derived from a distinct telomeric site and not from the previous sequences occupying the expression site.

The model shows that the different observations that have been obtained with different VSG genes are consistent with a common mechanism of activation for all VSG genes. Its central features are the distinction between the duplicative transposition promoting basic genes into the pool of expressible telomeric copies; a separate mechanism selecting among the telomeric genes; and the influence of stability, or rate of gene replacements, on the kind of observation that will be obtained with particular genes. Other models may be formulated, which do not involve the intermediate copy, but we believe they must share these features if they are to reconcile the disparity of observations with a common activation pathway for all VSG genes.

ACKNOWLEDGMENTS

We thank Dr E.N. Miller for performing immunofluorescence and Judith Creighton for technical assistance.

REFERENCES

1. Michels P, Bernards A, Van der Ploeg LHT, Borst P: Nucleic Acids Res 10:2353, 1982.
2. Van der Ploeg LHT, Liu AYC, Michels P, De Lange T, Borst P, Majmuder HK, Weber H, Veeneman GH, Van Boom J: Nucleic Acids Res 10:3591, 1982.
3. Young JR, Donelson JE, Majiwa PAO, Shapiro SZ, Williams RO: Nucleic Acids Res 10:803, 1982.
4. Donelson JE, Young JR, Dorfman D, Majiwa PAO, Williams RO: Nucleic Acids Res 10:6581, 1982.
5. Young JR, Shah JS, Matthyssens G, Williams RO: Cell 32:1149, 1983.
6. Borst P, Cross GAM: Cell 29:291, 1982.
7. Longacre S, Riaraud A, Hibner U, Buck G, Eisen H, Baltz T, Giroud C, Baltz D: Mol Cell Biol 3:410, 1983.
8. Laurent M, Pays E, Magnus E, Van Miervenne N, Matthyssens G, Williams RO, Steinert M: Nature 302:263, 1983.
9. Lumsden WHR, Herbert WJ: Trans R Soc Trop Med Hyg 69:205, 1975.
10. Miller EN, Turner MJ: Parasitology 82:63, 1981.
11. Rice-Ficht AC, Chen KK, Donelson JE: Nature 294:53, 1981.
12. Williams RO, Young JR, Majiwa PAO: Nature 299:417, 1982.
13. Turner MJ: Parasitol 21:69, 1982.
14. Turner MJ, Cordingley JS: In Carlisle MJ, Collins JF, Mosely BEB (eds): Cambridge: Cambridge University Press 1981, pp 313.
15. Majiwa PAO, Young JR, Englund PT, Shapiro SZ, Williams RO: Nature 297:514, 1982.