The switch region on *Leishmania major* chromosome 1 is not required for mitotic stability or gene expression, but appears to be essential

Pascal Dubessay, Christophe Ravel, Patrick Bastien, Lucien Crobu, Jean-Pierre Dedet, Michel Pagès and Christine Blaineau*

CNRS UMR5093 ‘Génomme et Biologie Moléculaire des Protozoaires Parasites’, Laboratoire de Parasitologie-Mycologie, Faculté de Médecine, 163 Rue Auguste Broussonet, F-34090 Montpellier, France

Received as resubmission June 10, 2002; Accepted July 17, 2002

ABSTRACT

The *Leishmania* genome project reference strain, *Leishmania major* Friedlin, is trisomic for chromosome 1. The complete sequence of this chromosome has revealed that the genes are grouped into two large clusters of the polycistronic type, each borne by one DNA strand and located on each side of a 1.6-kb sequence often termed the switch region. Several hypotheses concerning the role of this switch region have been put forward (region of initiation of transcription for both gene clusters, origin of replication or centromeric sequence). In the present study, we have deleted this region on the three copies of chromosome 1 by sequential targeted replacements. The absence of the switch region did not alter the mitotic stability of the three deleted chromosomes. This region therefore does not appear necessary for chromosomal replication or segregation. However, during the third targeting round, which aimed at knocking out the last switch region, a fourth copy of chromosome 1 that retained this region appeared in all clones analysed. This suggests that the persistence of this switch region is necessary for parasite survival. We then showed that the presence/absence of the switch region did not act upon the expression of a resistance marker gene inserted beforehand into the left gene cluster of the same chromosomal molecule. This result suggests that the presence of this 1.6-kb sequence is not necessary for the expression of all genes on chromosome 1.

INTRODUCTION

*Leishmania* are protozoal parasites of the family of Trypanosomatids responsible for significant worldwide human morbidity and mortality. The clinical features vary from mild cutaneous lesions to fatal mucocutaneous or visceral involvement depending on the causative species. The genome structure of this organism has now been well explored; 36 heterologous chromosomes, sized from ~0.3 to 3 Mb, are characterised for all Old World species, totalling ~35 Mb (±2 Mb) for a haploid genome (1,2). Although *Leishmania* is considered to be essentially diploid, aneuploidy (mainly in the form of trisomic chromosomes) has been observed in natural isolates; thus, the *Leishmania major* Friedlin strain (*LmF*), the reference strain for the genome sequencing project, possesses three homologues of chromosome 1 (2,3). From the complete sequence of this ~300-kb chromosome (GenBank accession no. AE001274), the three homologues appear to be almost identical in their coding regions; they are essentially distinguished by a size variation of ~30 kb in the right non-coding subtelomeric end between chromosomes 1a and 1b/c (3,4).

The most remarkable finding of the *Leishmania* and *Trypanosoma* genome sequencing projects certainly is the gene organisation in these protozoa. Most of the genes are grouped into large (50–300 kb) clusters of the polycistronic type grouped on the same DNA strand. These clusters can be transcriptionally divergent or convergent depending on the chromosome region (5). The structure of *LmF* chromosome 1 is one of the most conspicuous examples of this organisation; two clusters containing 29 and 50 coding DNA sequences (CDSs), each borne by one DNA strand and oriented toward one telomere, are joined by an A/T-rich 1.6-kb sequence, hereafter termed the switch region (4). Obviously, the particular position of this region as well as its structural features (e.g. A/T richness) (6) have lead to several hypotheses concerning its potential role, whether it be the site of transcription initiation for both clusters (4,7), a replication origin (5,8) or even a centromere (4). Notably, a negative correlation between purine excess and cumulative GC skew with (here termed divergent) coding directions suggested that the switch region might be the replication origin on this chromosome (8). Here, we report the successive deletions of this region on chromosome 1 homologues and analyse the effects of these targeted replacements upon mitotic stability or transcription.
MATERIALS AND METHODS

Cell line cultures and transfection conditions

The L. major Friedlin (LmF) strain (MHOM/IL/81/Friedlin) has been described previously (2). Promastigote culture and transfection experiments were as described previously (9), except for the selection of transfectant clones, which was done on RPMI1640 low-melting agarose plates (0.6%). Depending on the resistance gene(s) inserted, transfectants were selected with 30 μg/ml hygromycin B (Gibco Life Technologies®), 10 μg/ml phleomycin (Euromedex®) and/or 100 μg/ml puromycin (Euromedex®). After 10 days, resistant clones were isolated and cultivated for 1 week, and their genomic DNA was analysed by pulsed field gel electrophoresis (PFGE).

For the study of chromosomal stability, two clones, termed cHB1 and cHB2 (cHB for cΔ1.6::HYG/Δ1.6::BLE/Δ1.6::PAC/1.6), were cultivated without drug pressure, subcloning was carried out at different time intervals and the subclones cultivated in parallel in the presence or absence of drugs. For the study of transcription, clones cΔcH::HYG/Δ1.6::PAC 1, 2, 3 and 4 were selected with puromycin alone. These clones were then maintained in drug-free medium during 32 generations and then transferred cultivated in parallel in the presence or absence of drugs. Cell growth curves were established by daily cell counts on a Thoma haemocytometer after cultures were seeded with 2 × 10^5 cells ml⁻¹ at day 0.

Construction of the transfection vector

The transfection vectors used for this study were derived from a home-constructed vector termed pVV (GenBank accession no. AF315645), the construction of which has been detailed elsewhere (9). In this study, the HYG (10), BLE (11) and PAC (12) genes (conferring resistance to hygromycin B, phleomycin and puromycin, respectively) were used as selection markers. The flanking DNAs were from the 5' and 3' untranslated regions of the L. major dihydrofolate reductase-thymidylate synthase (DHFRTS) gene (13). Two homologous recombining sequences were used for all knock-outs of the switch region, and two others for the insertion of the HYG gene into the CCP2 gene on the left part of chromosome 1: their positions on chromosome 1 are indicated by the position of the primers used for their PCR amplification, referring to the complete sequence of chromosome 1 (GenBank accession no. AE001274): probe P1.6 (77384–79048) strictly encompasses the 1.6-kb switch region; probes PCCP (44837–45270), PORF20 (50717–51024), P3 (79559–80066) and P5 (171328–172306) correspond to segments of chromosome 1 located 33 and 26 kb to the left side, and immediately and 93 kb to the right side, of the switch region, respectively. Homologues of chromosome 1 were distinguished by probe ST515, specific for the 81-bp minisatelitte sequence LiSTIR1 (14) which is present in chromosome 1b/c and absent from the 1a homologue (3). The HYG, BLE and PAC probes were gel purified after electrophoresis of the corresponding enzyme-restricted pVV vector. The hybridisation conditions were as described by Ravel et al. (14).

Analysis of the recombination events

The replacements of the switch region on chromosome 1 by the drug resistance genes were analysed by double restriction with NheI and SpeI, followed by Southern analysis. NheI was selected for this analysis as its restriction sites encompass a ~9-kb fragment comprising the switch region. The SpeI site had been introduced by the insertion of the linearised transfection vector (see Fig. 2). The integration on the same chromosomal molecule of the HYG and PAC genes into CCP2 and the switch region, respectively, was verified by an SpeI restriction.

RESULTS

Targeted replacement of the switch region by drug resistance genes in the three copies of LmF chromosome 1

Three homologues of chromosome 1 are present in LmF, sized 285 and 315 kb for chromosome 1a and chromosome 1b/c, respectively (both 1b and 1c therefore being indistinguishable on a PFGE karyotype). Three rounds of transfection were carried out in order to knock out the 1.6-kb switch region in these three homologues. Transfectant clones were studied by PFGE and Southern analysis to verify the successful integration of the resistance genes. After the first transfection round, the HYG gene integrated on chromosome 1b/c in clone cΔ1.6::HYG, termed cH (Fig. 1), in replacement of the switch region locus (see below). A second transfection was realised on this clone and the BLE gene was then found integrated on chromosome 1a, with the switch region absent from this chromosome (clone cΔ1.6::HYG/Δ1.6::BLE, termed cHB; Fig. 1). Lastly, clone cHB was transfected in order to replace the last copy of the switch region present on the third homologue of chromosome 1 (b/c) by the PAC gene. Twenty-four clones (termed cHBP) expressing the three drug resistances were then isolated and analysed as above. Three interesting points were then found. (i) The switch region was still present in all 24 clones, on chromosome 1b/c. (ii) In 23 clones, the PAC gene integrated at the level of chromosome.
1a, but this did not knock out the BLE gene (see for example clone cHBP1 in Figure 1). (iii) In only one clone (cHBP2), PAC integrated at the level of chromosome 1b/c, but without knocking out HYG or the switch region in these homologues (Fig. 1). Two alternative hypotheses may explain these results: either two resistance genes had integrated the same chromosomal molecule (by a non-homologous recombination event) or the last transfection round had induced the formation of a fourth chromosome 1 homologue. We analysed the targeted replacements in more detail in order to settle this point.

Restriction analysis of the switch region replacements

The replacement of the switch region by the different resistance genes should introduce an SpeI site immediately upstream of the resistance gene (Fig. 2A). The correct replacements were thus verified by NheI–SpeI double digestion of total genomic DNA (see Materials and Methods). In clone cHBP1 (Fig. 2B), three NheI–SpeI fragments of the expected sizes (3812, 4043 and 4442 bp) were recognised by DNA probes of the three resistance genes, BLE, PAC and HYG, respectively. This result rules out the hypothesis of non-homologous recombination described above, and confirms the insertion of the three markers on three distinct chromosome 1 homologues. In addition, hybridisation with the switch region probe revealed an 8.7 kb fragment corresponding to the wild-type locus. The three remaining fragments (3.8, 4 and 4.4 kb in size) each hybridised with one of the drug resistance genes inserted (BLE, PAC and HYG, respectively); their respective size correlated with the expected size after replacement of the switch region by the resistance gene [X in (A)].

Figure 1. Targeted replacement of the 1.6-kb switch region by drug resistance genes in the three homologues of chromosome 1 in L.major Friedlin. (Top) The small chromosomes of the wild-type strain (W) and of different transfectant clones (cH, cHB, cHBP1 and cHBP2; see text for details) were separated on an ethidium bromide-stained PFGE gel. (Bottom) The same gel was Southern blotted and sequentially hybridised with DNA probes P1.6, HYG, BLE and PAC. After transfection of clone cHB by the PAC-containing vector, 24 clones were isolated. The 1.6-kb switch region was present at the level of chromosome 1b/c in all clones. In 23 of them (e.g., cHBP1 here), PAC integrated at the level of chromosome 1a, where BLE remained present. In only one clone (cHBP2), PAC integrated at the level of chromosome 1b/c where the switch region and the HYG gene are also present. No hybridisation was observed with all these probes outside of the chromosome 1 size class, in particular in the compression zone (not shown). Moreover, in all clones, chromosome 1-specific DNA probes CCP2 and P5 hybridised at the same level as chromosome 1a and 1b/c in the wild-type strain (not shown), showing that no major rearrangement occurred in these chromosomes.

Figure 2. Restriction analysis of the targeted replacements of the 1.6-kb switch region on chromosome 1 homologues. (A) Schematic representation of the NheI restriction fragment encompassing the switch region on the wild-type chromosome 1 (top line) and the mutated homologues (bottom line). In the first, the switch region (1.6) is comprised in an NheI fragment of ~8.7 kb. The insertion of any of the three drug resistance genes (R) used in our study strictly replaces the switch region and introduces an SpeI site. The size of R is different for each of the three resistance genes. Therefore, an NheI–SpeI double restriction should reveal a different sized fragment (X) depending on the specific resistance gene inserted. 5' and 3' are the L.major DHRFTS 5' and 3' untranslated regions used in the vector construct. The scale of the schematic representation is not proportional. (B) An NheI–SpeI double digest of clone cHBP1 DNA was electrophoresed on an 0.7% agarose gel and subsequently Southern blotted and hybridised. Hybridisation with DNA probe P3 [shown in (A)] revealed four different sized fragments corresponding to four homologues of chromosome 1. Probe P1.6 (specific for the switch region) hybridised with an 8.7 kb fragment corresponding to the wild-type locus. The three remaining fragments (3.8, 4 and 4.4 kb in size) each hybridised with one of the drug resistance genes inserted (BLE, PAC and HYG, respectively); their respective size correlated with the expected size after replacement of the switch region by the resistance gene [X in (A)].
on the other hand to the conserved wild-type locus. We concluded that four chromosome 1 homologues were present in clone cHBP1: two of the size of chromosome 1a (bearing either the BLE or the PAC gene) and two of the size of chromosome 1b/c (bearing either the HYG gene or the switch region). Similar restriction data were obtained in clone cHBP2, suggesting the presence of only one chromosome 1a and three homologues of the size of chromosome 1b/c (data not shown).

Mitotic stability of the switch region deletion mutant chromosomes

Clones cHBP1 and cHBP2 were cultivated without any drug pressure for 110 days (i.e. >130 generations in the conditions used here). Subcloning was carried out at different time intervals. The persistence of the resistance genes was directly tested by cultivating the subclones in parallel in the presence or absence of drugs. All the subclones tested conserved the three drug resistances. Moreover, the karyotypes of some of the subclones were verified by PFGE; all the subclones analysed exhibited a karyotype identical to that of the parental clone (not shown). Thus, the three mutant chromosomes bearing the switch region replacements proved mitotically stable in the absence of selective pressure. The presence of the switch region in cis therefore does not appear necessary to the mitotic stability of chromosome 1.

Role of the switch region in the transcription of chromosome 1

We then wished to further investigate the role of the switch region in the transcription of chromosome 1 gene clusters. In order to specifically identify the transcription of one of the homologues, in spite of the low nucleotide divergence between homologues in LmF (4), we inserted the HYG gene into the left gene cluster of chromosome 1a. The targeted sequence was the CCP2 gene, located at equal distance between the telomere and the switch region. HYG was inserted in the same orientation as CCP2. The selected clone (cΔccp::HYG) was then transfected with pVV-1.6/PAC, in order to replace the switch region by the PAC gene in a rightwards orientation (Fig. 3A). Four transfectants (cΔccp::HYG/Δ1.6::PAC) that had integrated PAC in chromosome 1a were isolated. They were selected and cultivated without hygromycin. Southern analysis of the molecular karyotype of these clones (Fig. 3B) showed the insertion of both resistance genes and the loss of the switch region all on chromosome 1a. SpeI digestion followed by hybridisation with a probe specific for ORF20 revealed the expected 34 kb (34 453 bp) fragment (not shown), confirming that both resistance genes were located on the same chromosomal molecule in the four clones. The latter as well as cΔccp::HYG were then cultivated in parallel in the presence or absence of hygromycin. All growth curves proved identical (Fig. 3C), showing that the HYG gene was expressed in a similar manner whether the switch region was present or absent.

DISCUSSION

In the present study, we have focused on the as yet unclear role of the 1.6-kb sequence located at the junction between the two large divergent gene clusters of L.major chromosome 1. We performed targeted deletions of this sequence on the three chromosome 1 homologues present in this strain. Nevertheless, we were unable to create null mutants for this switch region, as a fourth wild-type copy of chromosome 1

Figure 3. Analysis of clone cΔccp::HYG/Δ1.6::PAC bearing a double mutation on chromosome 1a. (A) Schematic map of the left part of chromosome 1a, including the two targeted sequences and the transfection vector sequences inserted; the arrows indicate the transcription sense orientation of the marker resistance genes. The wild-type chromosome 1 has only one SpeI site located at the right end of the chromosome (position 267 268 bp). The double insertion of HYG and PAC introduced two novel SpeI sites. The collinearity of both resistance genes on the chromosome was verified by the hybridisation of probe PORF20 on a ~34-kb SpeI fragment. (B) PFGE separation of the small chromosomes of strain LmF (a), clone cΔccp::HYG (b) and clone cΔccp::HYG/Δ1.6::PAC (c). The gel was Southern blotted and hybridised sequentially with the P1.6, HYG and PAC probes. (C) Comparison of the growth kinetics of clones cΔccp::HYG (a) and cΔccp::HYG/Δ1.6::PAC (b) and 2 (c) in hygromycin-containing medium over 7 days (J0–J7). Identical curves were obtained in drug-free medium.
persisted after the last transfection round, leading to tetrasomy for this chromosome. Other authors have experienced similar difficulties in disrupting essential genes in the *Leishmania* genome, due to intrachromosomal amplification, creation of extrachromosomal elements (16), genomic rearrangements of the duplication–transposition type (17), or duplication of whole chromosomes (18). We therefore conclude that the persistence of this 1.6-kb sequence is essential for parasite survival. In a second part of the study, we examined possible changes in the expression of the *HYG* gene inserted into the left gene cluster of chromosome 1a after we deleted the switch region on the same chromosome. The potential role of this region, as well as a number of mechanisms in which it is supposedly involved (replication, transcription, segregation), are discussed below.

The first hypothesis tested is whether this sequence is necessary for chromosomal stability, i.e. is a single replication origin (8) or a centromeric sequence (4). Our results show that its presence in *cis* is not necessary for chromosomal stability. Indeed, chromosomes which did not bear this sequence were mitotically stable, hence they were able to replicate and segregate normally. Regarding centromeric function, this finding is in agreement with recent studies suggesting that the sequences involved in chromosomal segregation may be located in the subtelomeric part of chromosome 1 (19). Regarding chromosome replication, although our results do not support the presence of an origin of replication in the switch region, they do not formally rule out this hypothesis. Indeed, in *Saccharomyces cerevisiae*, the knock-out of a functional autonomous replicating sequence (ARS) can be compensated by the activation of a neighbouring ARS (20,21). Moreover, a *trans*-complementation by an ARS present on the other homologue might be responsible for the maintenance of a chromosome fragment without origin (22). Here, the difficulty encountered in completely knocking out the switch region from the genome might be explained by the second hypothesis. In this case, the only remaining 1.6-kb region (on the fourth homologue) might complement the absence of origin on the three others. More elements will obviously be necessary to settle this point.

The second hypothesis that has been proposed about the role of the switch region is that it is a key element in the initiation/ regulation of transcription of the two gene clusters (4). The difficulty in knocking it out of the genome might be due to the fact that all essential genes present on chromosome 1 would remain untranscribed on the mutated homologues. However, our results do not fit with the latter hypothesis. First, the three resistance genes replacing the switch region in the knockouts were expressed. Secondly, when we deleted the switch region on a homologue of chromosome 1 tagged with the *HYG* gene integrated in the left gene cluster, the growth kinetics of the fourth homologue did not act upon the expression of this collinear gene.

The persistence of the switch region might then have another more trivial explanation: its knockout may have disrupted the splice site of the first gene in each of the opposing polycistronic transcription units [similar to exopolyphosphatase and poly(A) export protein genes, respectively]. If either of these genes were essential, one would never generate a switch region null mutant. Finally, one cannot rule out another more speculative hypothesis: that the switch region, which does not appear to contain any open reading frames, might contain essential non-coding RNAs (e.g. the recently identified miRNAs) (23).

On the whole, this work provides direct functional evidence that the main hypotheses about the role of the switch region (whether single origin of replication or region of initiation of the transcription of both chromosome 1 gene clusters) (4–8) are not validated by experimental results. The bioinformatic analysis of five switch regions located on different chromosomes did not allow identification of structural elements strictly common to all, with the exception of a high intrinsic DNA curvature (6). Although their peculiar position between two convergent or divergent gene clusters made them candidates as major regulating/functional sequences, it is not impossible that their role is restricted to the transcription of the sole adjacent genes. The publication of the complete sequence of the *LmF* genome, predicted for late 2002 (24), should allow a more systematic structural and functional analysis of these regions, and thus help to define their possible role in chromosome function.

**ACKNOWLEDGEMENTS**

We gratefully acknowledge the expert technical help of Michéle Lefebvre. We also thank Etienne Schwob for helpful advice and discussion. Finally, we are grateful to the reviewers for their pertinent questioning, which led us to more precise and extended conclusions. This study received financial support from the Centre National de la Recherche Scientifique (CNRS), the GDR 1077 of the CNRS/DGA-DSP and the CNRS programme ‘Puces à ADN’ (no. 00N65/0004).

**REFERENCES**


