MOLECULAR AND CELLULAR BIOLOGY, Dec. 1995, p. 6845–6853 Vol. 15, No. 12
0270-7306/95/$04.00+0

Increased Expression of *LD1* Genes Transcribed by RNA Polymerase I in *Leishmania donovani* as a Result of Duplication into the *rRNA* Gene Locus

MICHAEL J. LODES,[1]† GILLES MERLIN,[1]‡ THEO DEVOS,[1] ANIRBAN GHOSH,[2] RENTALA MADHUBALA,[2] PETER J. MYLER,[1,3] AND KENNETH STUART[1,3]*

Seattle Biomedical Research Institute, Seattle, Washington 98109-1651[1]; Pathobiology Department, University of Washington, Seattle, Washington 98195[3]; and School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India[2]

Received 13 March 1995/Returned for modification 18 May 1995/Accepted 8 September 1995

Eukaryotic protein-coding genes are generally transcribed by RNA polymerase II (Pol II), which has a lower transcription rate than that of Pol I. We report here the duplication of two *LD1* genes into the rRNA locus and their resultant transcription by Pol I. The multigenic *LD1* locus is present in a 2.2-Mb chromosome in all stocks of *Leishmania* spp. and is also present in multicopy 200- to 450-kb linear chromosomes or multicopy circular DNAs in over 15% of stocks examined. Genomic rearrangement in *Leishmania donovani* LSB-51.1 resulted in duplication of a 3.9-kb segment of *LD1* containing two genes (*orfF* and *orfG*) and of a 1.3-kb segment from ~10 kb downstream into the *rRNA* gene repeat region of the 1.2-Mb chromosome. Short sequences (12 or 13 bp) common to the 2.2-Mb *LD1* and 1.2-Mb *rRNA* loci suggest that this gene conversion occurred by homologous recombination. Transcription of the duplicated genes is α-amanitin resistant, indicating transcription by Pol I, in contrast to the α-amanitin-sensitive (Pol II) transcription of the genes in the 2.2-Mb *LD1* locus. This results in higher transcript abundance than expected from the gene copy number in LSB-51.1 and in elevated expression of at least the *orfF* gene product.

Several multigene loci undergo amplification in *Leishmania* spp. (for a review, see reference 3). These gene amplifications have often been the result of selection for drug resistance in vitro (4, 5, 21, 29, 37). However, unselected strains of *Leishmania* spp. also have amplified DNAs, including *LD1* (27, 52), but their significance is unknown. The amplicons are circular or linear molecules of various sizes and copy numbers, and the mechanism of their amplification appears to involve homologous recombination based on the presence of direct or inverted repeats at the termini of the amplified sequence. For example, rearrangement of the *H* region of *Leishmania tropica* has been shown to involve either direct or inverted DNA repeats of approximately 200 bp or greater (38, 55), while the circular *LD1* (*CD1*) of *Leishmania mexicana* M379 is thought to be generated by recombination involving short (13-bp) direct repeats (28).

We have been investigating the multigene *LD1* locus that is present within a 2.2-Mb chromosome in *Leishmania donovani* and a similar-sized chromosome in all *Leishmania* strains. *LD1* locus sequences are amplified as 20- to 60-copy, 200- to 450-kb chromosomes in some strains or 100- to 200-copy circular DNAs in others (52, 53). The *LD1* locus contains nine genes, including one (*orfB*) which encodes ribosomal protein L37 (35), one (*orfC*) with homology to genes which appear to have a role in regulation of cell growth (36), and another (*orfG*) with homology to *Trypanosoma brucei ESAG10* (18, 34). *ESAG10* and *orfG* predict proteins with 10 to 12 potential membrane-spanning domains, indicating possible transporter function (34).

In the Kinetoplastida, most protein-coding genes are transcribed as polycistronic units by RNA polymerase II (Pol II), similar to other eukaryotic systems. However, in *T. brucei*, the variant surface glycoprotein gene (*VSG*) and expression site-associated genes (*ESAG*s) are transcribed by Pol I (or a Pol I-like enzyme) in bloodstream forms, as are the procyclic acidic repetitive protein (*PARP*) and associated genes in procyclic forms (7, 11, 42). Translation of Pol I-transcribed protein-coding genes is possible in trypanosomatids because of *trans*-splicing of polycistronic pre-mRNAs (33, 48). The 39-nucleotide (nt) splice leader (or mini-exon) provides the 5′ cap necessary for message transport, stability, and translatability (22). Thus, in these organisms, the higher transcription rate of Pol I is used to ensure high levels of expression of these important genes.

We describe here the duplication of a 3.9-kb sequence containing *LD1 orfF* and *orfG*, along with a 1.3-kb sequence that is ~10 kb downstream of *orfG* in the 2.2-Mb chromosome, into the rRNA (*rRNA*) gene repeat locus of the 1.2-Mb chromosome. We find that the duplicated sequences are transcribed by Pol I, rather than Pol II which transcribes *LD1* genes within the 2.2-Mb chromosome, as well as amplified linear and circular DNAs. Thus, steady-state *orfF* and *orfG* transcripts are more abundant than one would predict given their gene copy number, and expression of at least the *orfF* gene product is substantially elevated. This genomic rearrangement appears to involve homologous recombination between short (12- or 13-bp) sequences in the *LD1* and *rRNA* loci. These junctions contain a 7-bp consensus sequence that also occurs at or near other recombination junctions of *LD1* sequences, suggesting possible involvement of specific sequence recognition in these recombination events.

* Corresponding author. Mailing address: Seattle Biomedical Research Institute, 4 Nickerson St., Seattle, WA 98109-1651. Phone: (206) 284-8846. Fax: (206) 284-0313. Electronic mail address: kstuart@u.washington.edu.

† Present address: Corixa Corporation, Seattle, WA 98104.

‡ Present address: URA CNRS 1637, Université Bordeaux 2, 33076 Bordeaux Cédex, France.

MATERIALS AND METHODS

Strains, culture conditions, and nucleic acids. Stocks used in this study are *L. donovani* MHOM/IN/80/DDB (LSB-8.1); *L. donovani* K42 (LSB-20.1); *L. donovani* BA147 (LSB-36.1); *L. donovani* MHOM/SD/00/Khartoum (LSB-51.1); *L. donovani* MHOM/SD/00/Khartoum LV711 (LSB-51.2); *L. donovani* MHOM/SD/00/1S-2D (LSB-52.1); *L. donovani* L13 (LSB-60.1); *L. mexicana* MNYC/BZ/62/M379 (LSB-13.1); and *Leishmania infantum* MHOM/BL/67/ITMAP 263 clone 10 (LSB-7.1), which were obtained as described previously (52). Promastigotes were cultured at 24°C in RPMI 1640 medium (Gibco BRL) supplemented with 10% heat-inactivated fetal calf serum (Gibco BRL) and 5% each of minimal essential medium (MEM) sodium pyruvate, MEM essential amino acids solution without L-glutamine, MEM nonessential amino acids solution, and glucose (0.2%). Total genomic DNA and agarose blocks containing 10^7 cells were prepared as previously described (31, 44). Total RNA was prepared by the guanidinium isothiocyanate method (10), and poly(A)$^+$ RNA was enriched with Poly(A) Quick columns (Stratagene), using supplied protocols and buffers.

Lambda library construction and clones. A lambda library was prepared from LSB-51.1 total genomic DNA with a Lambda FIXII/Gigapack Gold Cloning Kit following the manufacturer's (Stratagene) instructions. Briefly, total genomic DNA was partially digested with the enzyme *Sau*3AI (Gibco BRL), using limiting enzyme; fragments in the 10- to 25-kb size range were isolated on a sucrose gradient; fragment ends were partially filled with dATP and dGTP, ligated to λ arms, and packaged. Clones λ5, λ7, λ15, and λ19, representing the appropriate regions of the *LD1* sequence from the 2.2-Mb chromosome, were obtained by screening the library with mixed hexamer probes, derived from clones pS21 (λ7) and pSK12 (λ5 and λ19) (34, 53) (see Fig. 4 and 6 for locations of clones and subclones). Clones λ1AA and λ2AA, representing the duplicated copy of *FG*x on a 1.2-Mb chromosome, were obtained by first screening the LSB-51.1 λ library with a probe for *orfF* (pEK1) (34), purifying positive plaques, and rescreening with a probe for the *T. brucei 18S rRNA* gene (pGH174) (23). Regions of the λ clones were subcloned into pBluescript II SK$^-$ (Stratagene) by standard techniques as follows: clone λ5 as a 4.2-kb *Hin*dIII-*Not*I fragment (pHN42) containing *LD1 orfD*, *orfE*, and *orfF* and as a 2.6-kb *Bam*HI-*Sal*I fragment (pBS26) containing *orfG*; clone λ7 as a 3.5-kb *Not*I-*Hin*dIII fragment (pNH3.5) containing the 5′ end of the *LD1* sequence and flanking sequence; λ15 as a 3.5-kb *Bam*HI fragment (pB3FL) from which was derived a 1.1-kb *Sal*I fragment (p11S) containing the 3′ end of the *LD1* sequence and flanking sequence; and λ19 as a 3.4-kb *Bam*HI fragment (pB28) containing the duplicated *LD1* flanking region, sequence x. This region was also cloned as a 2.2-kb *Bam*HI-*Xba*I fragment obtained from PCR amplification of clone λ19 DNA (pBX19) and from total LSB-51.1 DNA (pBX511) with primers 242-93 (ATGTCTAGAAGTCCGTA CAGGTACG) and 303-93 (GTGCGCTGCTCTCGTTTCC). The *rRNA* gene external spacer region was subcloned as a 4.5-kb *Eco*RI fragment (pE45) from clone λ2AA. The 1.9-kb *Not*I-*Eco*RI fragment (pNE19) containing duplicated *orfF*, the 1.1-kb *Eco*RI-*Not*I fragment (pEN11) containing duplicated *orfF* and *orfG*, and the 7.8-kb *Not*I-*Eco*RI fragment (pNE78) containing duplicated *orfG*, x sequence, rRNA gene external spacer, promoter, and *18S rRNA* gene were obtained from λ1AA. In addition, 0.9-kb *Eco*RI-*Not*I clones (1AA09, 51109, and 51209) containing the *18S rRNA* gene precursor sequence were obtained by PCR amplification of clone λ1AA DNA and total genomic DNAs from strains LSB-51.1 and LSB-51.2, respectively, with the primers 327-94 (AGAATTCGGCAT GCATGGCTAAGTCC) and 328-94 (GAAGGGGAGCGGCCGCGGTGCG TGGATGCCG). A cDNA library was prepared from LSB-51.1 poly(A)$^+$-enriched RNA by using a Uni-ZAP XR/Gigapack II Gold Cloning Kit following the manufacturer's (Stratagene) instructions. cDNA clones (cF9 and cG47) were isolated by screening the library with probes obtained by labeling inserts from clones pEK1 and pEB1, respectively, with [α-^{32}P]dATP by using mixed hexamer priming (Pharmacia). Additional clones used in this study, including pT15, pE13, pH30, pCB1, pBH11, pH8, pNS1, pS22, pSN1, pNH1, pHC1, pCK1, pEK1, pEB1, pB22, pS5, pS10, pSK1, and pK8, are subclones of previously described clones (34, 53) (see Fig. 6). A 3′ *Not*I fragment (NKI) from subclone pK27 (53) was used as the template for mixed hexamer priming. A clone containing *T. brucei* α- and β-tubulin genes (pTbαβT-1) has been described previously (50).

Nuclear runoff analysis. Isolation of nuclei, labeling of nascent transcripts, and hybridization to filters were done as described previously (19, 32). Briefly, nuclei from cultured *L. donovani* promastigotes, strains LSB-51.1, LSB-51.2, LSB-8.1, and LSB-7.1 (10^9 cells per assay), were isolated by Dounce homogenization of cells in modified lysis buffer B (10 mM Tris-Cl [pH 7.4], 3 mM $CaCl_2$, 2 mM $MgCl_2$, 0.5% Nonidet P-40) and stored in 40% glycerol at −70°C. Transcripts were labeled with [α-^{32}P]UTP in the presence or absence of 1 mg of α-amanitin per ml and isolated, after DNase I (Gibco BRL) and proteinase K (Sigma) digestion, by trichloroacetic acid precipitation. Filters were prepared for hybridization by linearizing 5 μg each of plasmid DNA with the appropriate enzyme, denaturing with 0.1 M NaOH, neutralizing with 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and transferring to Nytran filters with a slot blot apparatus (Schleicher and Schuell).

Nucleic acid separation and transfer. Chromosome-sized DNA was separated by clamped homogeneous electric field gel electrophoresis with the BioRad Mapper system in 0.8% SeaKem LE agarose (FMC) gels and 1× TAE buffer using a linear switch time gradient of 1 min and 26.25 s to 2 min and 41.34 s at 6 V/cm for 24 h at 14°C. Digested genomic DNA was separated by field inversion gel electrophoresis in 0.8% LE agarose (FMC) and 1× TBE (Tris-borate-EDTA) buffer with interval ramping of 0.5 to 1.5 s at 7 (cathode) and 4.5 (anode) V/cm for 24 h at 14°C. Gels were stained with ethidium bromide and transferred to Nytran filters (MSI) by capillary action (43). Total RNA (10 μg) and poly(A)$^+$-enriched RNA (2 μg) were separated in either 1.0, 1.25, or 1.5% SeaKem ME agarose gels containing 2.2 M formaldehyde and 20 mM MOPS (morpholinepropanesulfonic acid), pH 7.0, and transferred as previously described (15). DNA and RNA were cross-linked to Nytran filters by UV irradiation following the manufacturer's instructions (Stratalinker UV crosslinker; Stratagene). Filters were prehybridized, hybridized, and washed as previously described (30). [α-^{32}P]UTP-labeled riboprobes were prepared by transcribing gel-isolated restriction-enzyme-digested fragments with T3 or T7 RNA polymerase (Stratagene). Unincorporated label was removed with a Bio-Gel P-30 (Bio-Rad) column. DNA was labeled with [α-^{32}P]dATP by mixed hexamer priming (Pharmacia) of gel-isolated DNA fragments by using the Klenow fragment of DNA polymerase I.

RNase H cleavage. RNase H digestion was performed as previously described (9). Briefly, 10 μg of total RNA from LSB-51.1 and LSB-20.1 was mixed with 200 ng of oligo(dT)$_{12–18}$ (Pharmacia) and 120 ng of a second oligonucleotide, 223-92 (TCACCAGCCACGGCGCATTG) (designed to cleave the *orfG* transcript as shown in Fig. 1B), and incubated at 65°C for 2 min and then at 42°C for 15 min. Samples were then digested with 3 U of RNase H at 37°C for 30 min, ethanol precipitated, separated on 1.25% formaldehyde agarose gels, transferred to Nytran filters, and hybridized with the appropriate antisense riboprobe.

Sequencing and sequence analysis. Double-stranded DNA was sequenced with a Taq Dye Primer or Dye Terminator Cycle Sequencing kit (Applied Biosystems, Inc.) and a model 373A DNA sequencer (Applied Biosystems, Inc.). Computer analysis of nucleotide sequence was performed with DNASTAR (DNASTAR Inc., Madison, Wis.), ESEE (8) and University of Wisconsin Genetics Computer Group (12) software.

Detection of *LD1* gene products. The shorter *orfF* open reading frame (34) was PCR amplified from clone pB30 with oligonucleotide primers 5BorfF (TCGGATCCAATGCAAAGCGACGCACG) and 3EorfF (CCGCTCGAGCTCGC CGTTGGTGATGAGCTGCA). The PCR product was digested with *Xho*I and *Bam*HI, ligated with *Xho*I-*Bam*HI-digested pET17b (Novagen), and transformed into *Escherichia coli* DH5α cells (Gibco BRL) to obtain clone pET17bF-C14. Plasmid DNA was prepared and used to transform *E. coli* BL21 cells (Novagen), and expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) according to the manufacturer's instructions. The inducible recombinant protein of the predicted size (34 kDa) was isolated by dissolving inclusion bodies in 8 M urea; precipitation with $(NH_4)_2SO_4$ after dialysis to remove urea, and separation by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Gel slices containing the recombinant *orfF* protein were ground up, lyophilized, resuspended in 10 mM Tris (pH 8.0), and ~1 mg of protein was mixed with an equal volume of incomplete Freund's adjuvant and 100 μg of *N*-acetylmuramyl-L-alanyl-D-isoglutamine. A New Zealand White rabbit was immunized with 300 μg of recombinant in this adjuvant mixture and given booster doses of 150 to 250 μg of protein seven times at ~2-week intervals before sacrifice to obtain antiserum.

Promastigote cells from several strains of *Leishmania* spp. were harvested at mid-late log phase, spun down, washed, and resuspended to 5×10^9 cells per ml in SDS-PAGE buffer (0.15 M Tris [pH 6.8], 4.0% SDS, 20% glycerol, 1.4 M 2-mercaptoethanol, 0.06% bromophenol blue). Samples (10 μl) representing 2×10^7 to 5×10^7 cells were boiled before separation on an SDS–10% polyacrylamide gel and proteins were transferred to nitrocellulose filters by Western blotting (51). Filters were washed three times for 5 min each with phosphate-buffered saline (PBS) plus 0.4% Tween 20 and blocked by incubation with 5% nonfat milk in PBS for 30 min at room temperature. After the filters were washed (three times for 10 min each) in PBS plus 0.1% Tween 20, they were incubated for 60 min at room temperature with a 1/200 dilution of rabbit antiserum in the same buffer and washed again as described above. They were then incubated with a 1/20,000 dilution of ^{125}I-labeled protein A (ICN Biomedicals) for 60 min at room temperature. The filters were washed (twice for 1 min each and then three times for 10 min each) in PBS plus 0.1% Tween 20 and then washed for 5 min in a solution of PBS, 0.1% Tween 20, and 0.05% SDS. Bound antibodies were then visualized by autoradiography.

Nucleotide sequence accession numbers. GenBank accession numbers for the duplicated *FG*x sequence, sequence x, and the *18S rRNA* precursor and promoter sequences are L38571, L38570, and L38572, respectively.

RESULTS

Transcription of *LD1* genes. We examined *LD1* gene (*orfA-orfI*) transcripts (34–36) in several strains of *L. donovani*, which contain different *LD1* amplicons (circular molecules and large or small linear chromosomes) with a range (2 to 100) of copy numbers. Northern (RNA) blot analysis showed that, in general, steady-state transcript abundance was correlated with gene copy number (34, 36). This is illustrated in Fig. 1A, where

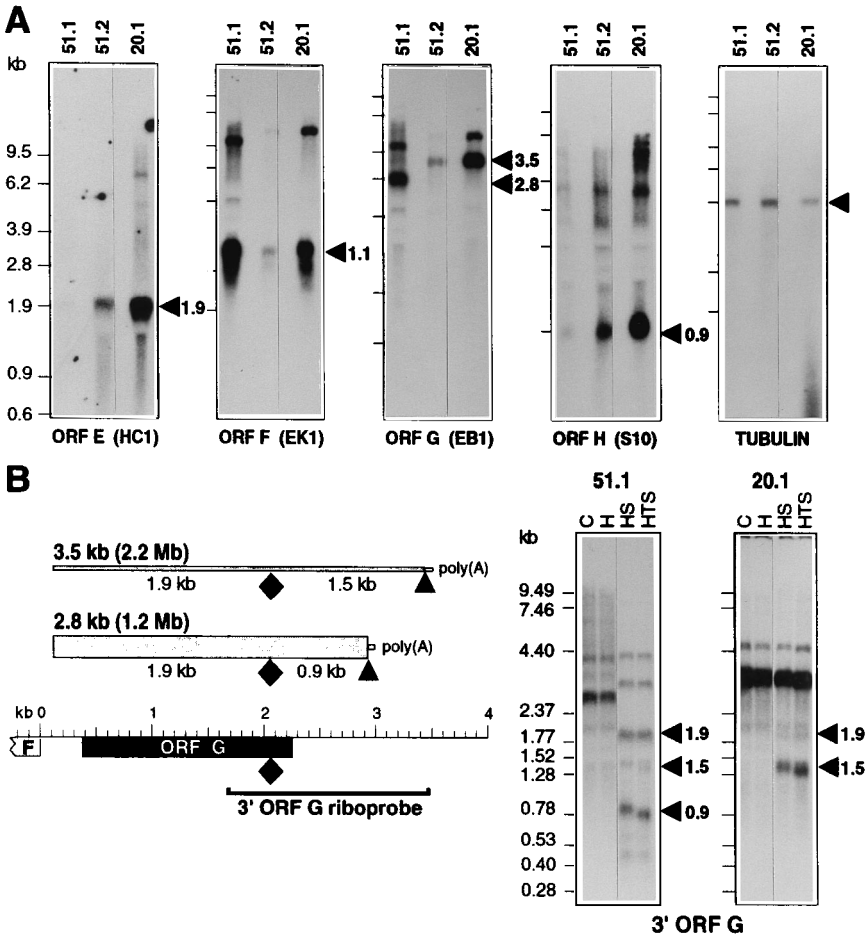

FIG. 1. Increased *orfF* and *orfG* transcript abundance in strain LSB-51.1 is not due to differential polyadenylation. (A) Samples (10 μg each) of total RNAs from LSB-51.1, LSB-51.2, and LSB-20.1 (shown as 51.1, 51.2, and 20.1, respectively, over the lanes) were separated on denaturing agarose gels, transferred, and hybridized with riboprobes specific for *LD1 orfE*, *orfF*, *orfG*, and *orfH* (clones pHC1, pEK1, pEB1, and pS10, respectively [shown as HC1, EK1, EB1, and S10, respectively, under the gels]), as well as a mixed hexamer-primed tubulin probe (pTbαβT-1). Major transcripts (and their sizes [in kilobases]) detected by each probe are indicated by arrowheads to the right of each gel. Promega RNA markers are shown by the tick marks to the left of each gel. (B) Samples (10 μg each of total RNAs from LSB-51.1 and LSB-20.1 were separated on denaturing agarose gels after no treatment (lanes C), RNase H digestion alone (lanes H), hybridization with oligonucleotide 223-92 followed by digestion with RNase H (lanes HS), or hybridization with oligonucleotides dT_{12-18} and 223-92 followed by digestion with RNase H (lanes HTS). The RNAs were transferred to Nytran filters and hybridized with a 3′ *orfG* riboprobe which was synthesized from *Kpn*I-digested clone pS85. The approximate size, abundance (indicated by thickness), and genomic origin of the LSB-51.1 *orfG* transcripts are shown diagramatically to the left of the gels. The RNase H cleavage sites using oligonucleotides dT_{12-18} and 223-92 are indicated by triangles and diamonds, respectively, and the sizes (in kilobases) of cleaved transcript fragments, which are also shown in the diagram, are indicated by arrowheads to the right of each gel. Gibco BRL RNA markers are indicated by the tick marks to the left of each gel.

the ~1.9-kb *orfE* and ~0.9-kb *orfH* mRNAs are least abundant in strain LSB-51.1, which contains only 2 (diploid) copies of these genes, and increase in abundance in LSB-51.2, which contains an additional 10 copies on a 450-kb minichromosome (52), and LSB-20.1, which contains ~200 copies on a 54.4-kb circular molecule (52). However, *orfF* and *orfG* mRNA levels in LSB-51.1 are much higher than those for the flanking *orfE* and *orfH* genes; indeed their abundance is even greater in this isolate than in LSB-20.1 (Fig. 1A). In addition, while the *orfF* transcript was unchanged in size (~1.1 kb) in all strains tested, an abundant ~2.8-kb *orfG* transcript was seen in LSB-51.1, as well as the less-abundant ~3.5-kb transcript seen in all strains (see arrowheads in Fig. 1A). An ~3.9-kb transcript was also detected in LSB-51.1 with both *orfF* and *orfG* probes, in addition to the ~4.6-kb transcript seen in all isolates, suggesting that these are precursors of the mature *orfF* and *orfG* mRNAs. Northern blot analysis with poly(A)$^+$ RNA showed that the mature *orfG* transcripts were polyadenylated (data not shown) and RNase H digestion using oligo$(dT)_{12-18}$ and an internal oligonucleotide revealed that the poly(A) tail length was ~60 nt for both *orfG* transcripts in LSB-51.1 (Fig. 1B). Similar results were obtained for other *LD1* mRNAs, including *orfF*, in LSB-51.1 and other strains (data not shown). Thus, the increase in *orfF* and *orfG* transcript abundance and change in size of the *orfG* transcript in LSB-51.1 is unlikely to result from differences in poly(A) tail length.

The 5′ and 3′ ends of the ~2.8- and ~3.5-kb *orfG* mRNAs were localized by RNase H digestion after hybridization with an *orfG*-specific oligonucleotide (see diagram in Fig. 1B). The ~3.5-kb *orfG* mRNA (in LSB-20.1 and LSB-51.1) is cleaved into an ~1.9-kb 5′ fragment and an ~1.5-kb 3′ fragment (see arrowheads in Fig. 1B). In LSB-51.1, cleavage of the ~2.8-kb *orfG* transcript results in the same ~1.9-kb 5′ fragment and an ~0.9-kb 3′ fragment. Thus, the 5′ ends of the ~3.5- and ~2.8-kb *orfG* transcripts appear to be the same, while the 3′ end of the ~2.8-kb, LSB-51.1-specific transcript is ~0.7 kb shorter than that of the ~3.5-kb transcript, suggesting the two transcripts utilize different polyadenylation sites.

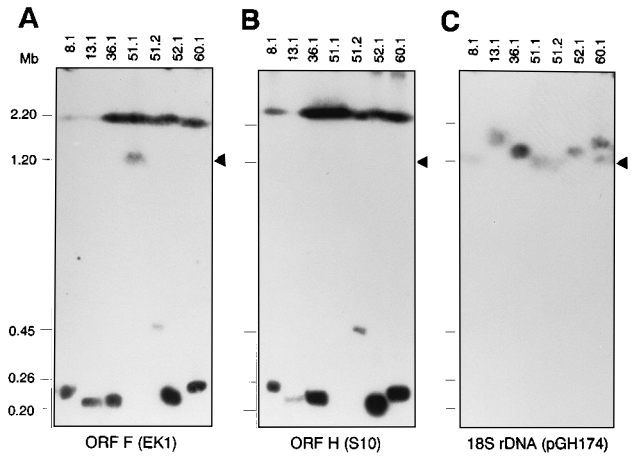

FIG. 2. Duplicated *orfF* and *orfG* genes are located on a 1.2-Mb chromosome in strain LSB-51.1. Chromosomal DNAs from several strains (strain designations over the lanes are minus LSB-) of *L. donovani* were separated by clamped heterogeneous electric field gel electrophoresis, transferred, and hybridized with mixed hexamer-primed probes for *orfF* (pEK1 [EK1]) (A), *orfH* (pS10 [S10]) (B), and *18S rRNA* (pGH174) (C). The arrowheads indicate the position of the 1.2-Mb chromosome which hybridizes only with probes for *orfF* and *orfG* and the *18S rRNA* gene in LSB-51.1. Chromosome size markers, shown by tick marks to the left of each gel, are derived from *Saccharomyces cerevisiae* and *Hansenula wingei* standards (Bio-Rad).

Rearrangements of the *LD1* locus. Since amplification of the *LD1* locus resulting in increased transcript abundance was frequently observed in *Leishmania* spp. (34, 36, 52), we examined the chromosomal distribution of *LD1* sequences in strain LSB-51.1. Southern blot analysis of clamped homogeneous electric field gel-separated chromosomes showed hybridization of an *orfF* probe to the 2.2-Mb chromosome in all strains and a 1.2-Mb chromosome in LSB-51.1 alone (Fig. 2A). Similar results were obtained with an *orfG* probe (data not shown), but probes for 5′ (*orfE*) and 3′ (*orfH*) flanking genes did not hybridize to the 1.2-Mb chromosome (data not shown and Fig. 2B). These results suggest that in LSB-51.1, a segment of the *LD1* locus containing *orfF* and *orfG* was duplicated into the 1.2-Mb chromosome.

Hybridization with an *18S rRNA* probe indicated that the 1.2-Mb chromosome also contains the *rRNA* gene locus (Fig. 2C), so we investigated the possibility that the duplicated *LD1* sequences were adjacent to the rRNA genes. Southern blot analysis of *Bam*HI-digested genomic DNA with an *orfG* probe showed hybridization to an ~5.2-kb fragment in all strains (Fig. 3A), indicating that it is derived from the *LD1* locus on the 2.2-Mb chromosome (which is present in all strains). In LSB-51.1 alone, the *orfG* probe hybridized with an additional ~29-kb *Bam*HI fragment, which must therefore be derived from the 1.2-Mb chromosome. The *18S rRNA* probe hybridized to an ~100-kb fragment in all strains, indicating that it is derived from the unaltered *rRNA* locus on a 1.2-Mb chromosome. In addition, the ~29-kb *Bam*HI fragment in LSB-51.1 also hybridized with the *18S rRNA* probe, suggesting that this fragment contained both *orfG* and *18S rRNA* sequences in LSB-51.1. Similar results were obtained with *Not*I- and *Eco*RI-digested DNA (Fig. 3B and C), indicating that duplication of the *orfF* and *orfG* genes (*FG* duplication) occurred within or near the *rRNA* gene locus. The exact location of the *FG* duplication was determined from restriction mapping and partial sequence analysis of genomic λ clones from this region (Fig. 4). Comparison of the *orfF* and *orfG* sequence from these clones with that obtained from the corresponding region of λ clones

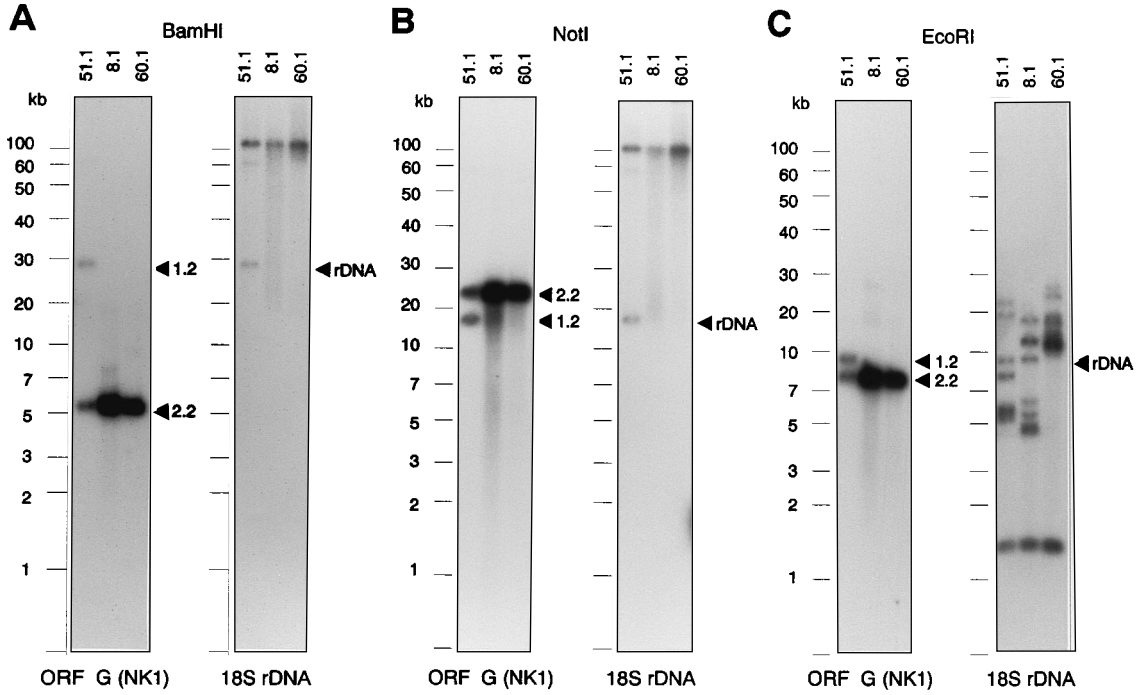

FIG. 3. The *FG* duplication is linked to the *rRNA* gene locus. Total genomic DNAs from three strains of *L. donovani* were digested with *Bam*HI (A), *Not*I (B), and *Eco*RI (C), separated by field inversion gel electrophoresis, and transferred, and the duplicate filters were hybridized with mixed hexamer-primed probes for *orfG* (NK1) or *18S rRNA* (pGH174). Arrowheads indicate the restriction fragments predicted for *LD1* from the 2.2-Mb chromosome (2.2), the *FG*x duplication on the 1.2-Mb chromosome (1.2), and the LSB-51.1-specific rRNA genes (rDNA). Size markers, shown at the left, are derived from Gibco BRL 1-kb and Bio-Rad 5-kb ladders. Strain designations over the lanes are minus LSB-.

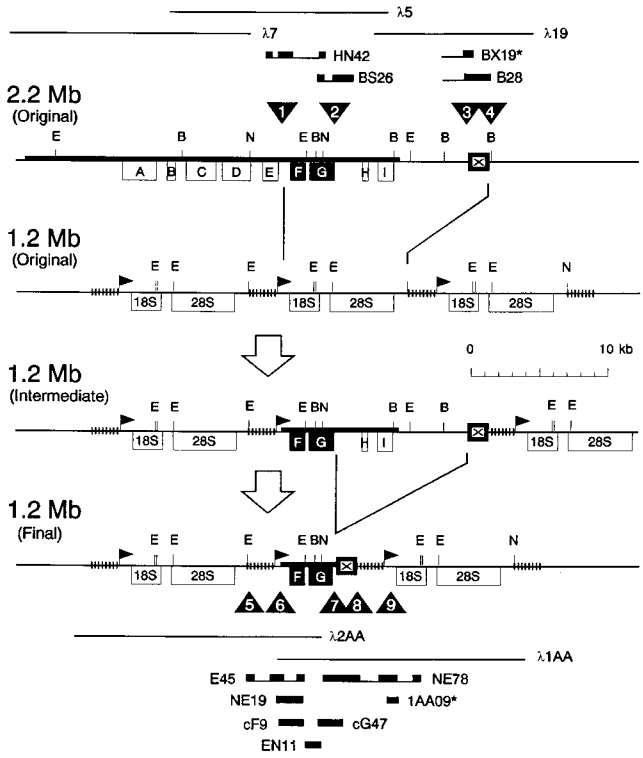

FIG. 4. Physical maps of the *LD1* locus on the 2.2-Mb chromosome, the *rRNA* gene locus, and *FG*x duplication on the 1.2-Mb chromosome. The origins of genomic λ clones and subclones, cDNA clones, and clones derived by PCR amplification are shown by lines above and below the appropriate genomic map. Plasmid designations in figure are shown minus the initial lowercase p. Asterisks indicate that several independent PCR clones were obtained for these regions. Sequenced regions of clones are indicated by thick lines within respective clones. *LD1* genes (*orfA-orfI*), as well as the *18S* and *28S rRNA* genes are boxed, and the duplicated *LD1* flanking sequence (x) is indicated by the thick box. The thick line within the genomic map indicates the 27.2-kb *LD1* sequence. Flags denote the inferred locations of *rRNA* Pol I promoters, and vertical bars indicate the region of 64-bp repeats in the *rRNA* gene external spacers. Triangles (numbered 1 to 9) indicate the locations of short sequence homologies (see Fig. 5) which are involved in the recombination event(s) denoted by vertical lines between maps. The arrows indicate the putative sequence of events resulting in the duplication of the ~15-kb *LD1* sequence into the *rRNA* gene locus and subsequent (or concomitant) deletion of a ~10-kb segment of this duplication. Restriction enzyme abbreviations: B, *Bam*HI; E, *Eco*RI; N, *Not*I.

derived from the *LD1* locus of the 2.2-Mb chromosome of LSB-51.1 revealed that the 5′ and 3′ boundaries of the *FG* duplication were 666 bp upstream of the *orfF* spliced leader site (740 bp upstream of the *orfF* ATG) and 40 bp downstream of the *orfG* stop codon, respectively.

The 5′ boundary of the duplicated *FG* sequence was ~360 bp downstream of the putative *rRNA* promoter. The location of the *rRNA* promoter is inferred from primer extension studies mapping the 5′ end of the rRNA precursor (~1.1 kb upstream of the *18S rRNA* gene [29a]) and by homology of this region to that from *Crithidia fasciculata* (20). The sequence immediately downstream of the *FG* duplication did not match that from the *rRNA* gene locus. Hybridization analyses (data not shown) using a probe from this region indicated that it was derived from a sequence ~10 kb downstream of *orfG* in the *LD1* locus (Fig. 4). Sequence analysis of genomic λ clones and cloned PCR fragments from this region (Fig. 4) identified the 5′ and 3′ boundaries of this 1.3-kb duplicated fragment (Fig. 4, sequence x) by comparison with sequence from the *FG* region and *rRNA* gene locus. Examination of sequence from this region and its immediate 5′ flank in the 2.2-Mb chromosome (Fig. 4) indicated a potential protein-coding open reading frame (*orf*γ) and Northern blot analysis revealed an ~2.4-kb transcript from this region in all strains (unpublished data). Analysis of cDNA clones from LSB-51.1 revealed that the duplication of sequence x downstream of the duplicated *orfG* sequence in the 1.2-Mb chromosome replaced the normal (2.2-Mb) *orfG* poly(A) site with that derived from *orf*γ (in sequence x), resulting in a shorter (2.8-kb) transcript from the duplicated *orfG* (Fig. 1). Comparison with sequence from the *rRNA* locus revealed that the 3′ boundary of the *FG*x duplication occurred in the 64-bp repeats within the *rRNA* nontranscribed spacer region (i.e., downstream of the *rRNA* genes and upstream of the promoter for the next *rRNA* repeat).

The *LD1* rearrangement involves homologous recombination. Alignment of sequence from the boundaries of the *FG*x duplication on the 1.2-Mb chromosome with the corresponding sequences from the *LD1* locus within the 2.2-Mb chromosome and the *rRNA* gene locus within the 1.2-Mb chromosome revealed short regions of homology (Fig. 5). A 13-bp sequence upstream of *orfF* is repeated downstream of the putative *rRNA* promoter and a 12-bp sequence at the 3′ boundary of sequence x is repeated 3′ to the *28S rRNA* gene. A 13-bp sequence is repeated 3′ to *orfG* and at the 5′ end of sequence x. Thus, it appears that homologous recombination between the 2.2-Mb *LD1* locus and the 1.2-Mb *rRNA* gene locus led to replacement of one or more repeats of the *rRNA* genes with the *FG*x sequence. Concomitant or subsequent recombination between the 3′ *orfG* and 5′ x sequences led to deletion of the ~10-kb intervening sequence (Fig. 4). Examination of the sequences surrounding these recombination sites revealed a 7-bp consensus sequence (tgtGCat) within or near all six regions of homology (Fig. 5). This sequence is also found at or near the boundaries of amplified *LD1* sequence in three additional strains (unpublished data), suggesting a possible role in the recombination-amplification process.

The duplicated *LD1 orfF* and *orfG* genes are transcribed by Pol I. The high levels of mRNA from *orfF* and *orfG* in the duplicated *FG*x region of the 1.2-Mb chromosome and their locations downstream of the putative *rRNA* promoter (Fig. 4) suggested the possibility that they are transcribed by Pol I, rather than Pol II, which is normally used for transcription of protein-coding genes. We utilized nuclear run-on transcription studies in the presence or absence of 1 mg of α-amanitin per ml to investigate this possibility. In the absence of α-amanitin, ^{32}P-labeled nascent transcripts from all strains tested (LSB-51.1, LSB-51.2, LSB-8.1, and LSB-7.1) hybridized to cloned DNA covering the entire *LD1* sequence, as well as an *18S rRNA* clone, and a cloned tubulin gene control (Fig. 6B and C, CONTROL panels). In general, the hybridization signal to *LD1* genes was higher for LSB-51.2 than for LSB-51.1, reflecting the increased *LD1* copy number (~10 copies on a 450-kb minichromosome) in the former. This is particularly apparent for longer clones such as pNH3.5, pH30, and p11S. The exceptions are clones pCK1, pEK1, and pEB1 which show higher signal than other clones in LSB-51.1, reflecting increased transcription of the duplicated *orfF* and *orfG* genes from the *rRNA* locus. Hybridization to clone pB22 is also elevated in LSB-51.1, relative to other clones in this strain, but since only a portion of the clone is represented in the *FG*x duplication, the signal is not as high as in LSB-51.2. When nascent transcripts from strain LSB-51.2 (Fig. 6C) or strains LSB-8.1 and LSB-7.1 (data not shown) were labeled in the presence of α-amanitin, ≥90% of the hybridization signal was eliminated for all clones except the *18S rRNA* clone. These data indicate the *LD1* genes in the 2.2-Mb chromosome (all strains), and small linear (LSB-51.2

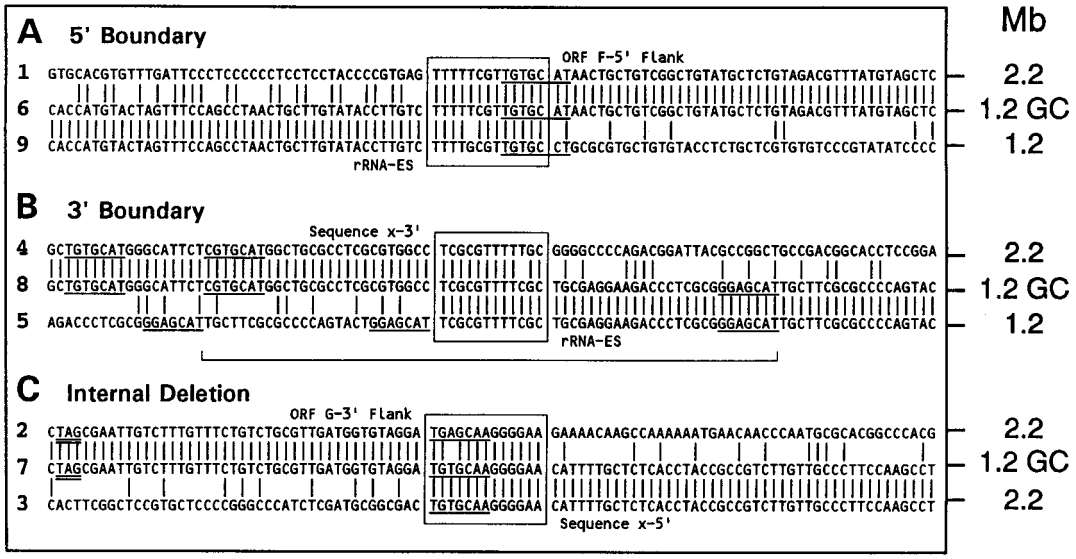

FIG. 5. Homologous recombination at the boundaries of *FG*x duplication and deletion. Sequences from clones representing the boundaries of the duplication-deletion event(s) on the 1.2-Mb chromosome in strain LSB-51.1 (indicated by 1.2 GC to the right of the sequence) were compared with sequences from the corresponding regions of the 2.2-Mb chromosome (2.2) and the *rRNA* gene locus (1.2). Labels above or below sequences indicate the region of origin. Sequences shown are from the 5′ (A) and 3′ (B) boundaries of the duplication and the boundaries of the internal deletion event (C). Regions of homology among aligned sequences are boxed, a 7-nt consensus sequence (tgtGCat) is underlined, and the *orfG* stop codon (sequences 2 and 7, panel C) is double underlined. A single 64-bp repeat unit from the *rRNA* external spacer (rRNA-ES) is demarcated by a bar below sequence 5 in panel B. ORF, open reading frame.

and LSB-8.1) or circular (LSB-7.1) amplicons are transcribed by Pol II (which is sensitive to α-amanitin). In contrast, α-amanitin treatment of LSB-51.1 (Fig. 6B) did not diminish hybridization to clones containing *orfF* or *orfG* (indicated by brackets), although the hybridization signal to all other clones (except the *18S rRNA* clone) was diminished to the same extent as in other strains. Thus, the duplicated *orfF* and *orfG* genes in the 1.2-Mb chromosome of LSB-51.1 are transcribed by Pol I (which is resistant to α-amanitin), presumably from the upstream *rRNA* promoter.

***LD1 orfF* and *orfG* genes are overexpressed in LSB-51.1.** Duplication of *orfF* and *orfG* into the *rRNA* locus and their consequent transcription by Pol I results in elevated mRNA levels relative to those of other *LD1* genes in strain LSB-51.1 (Fig. 1). Comparison of the duplicated *FG*x sequence (accession number L38571) with that from the corresponding region of the 54.4-kb circular *LD1* from LSB-7.1 (L25643) (34) indicates only two single-nucleotide differences. One of these (an A at position 2764 in L38571 versus a G in L25643) causes single amino acid difference in the predicted *orfG* protein sequence (Y in place of C). The other (loss of a C residue at position 1419 in L38571) results in the shortening of the open reading frame predicted from *orfF*. However, since as previously noted (34), the spliced leader site for *orfF* lies downstream of this region and thus a more 3′ ATG is likely used for the initiation of translation, this should cause no difference in the protein sequence of the *orfF* gene product. It is interesting to note that the sequence of at least one allele from the 2.2-Mb *LD1* locus in LSB-51.1 matches the sequence from the 54.4-kb circular *LD1* (nine C residues) rather than that from the duplicated *FG*x sequence (eight C residues) (data not shown). Thus, neither of the duplicated *orfF* and *orfG* genes in the *rRNA* locus of LSB-51.1 appear to be pseudogenes.

Western blot analysis using polyclonal antisera raised against recombinant *orfF* protein detects a protein of the size (~32 kDa) predicted from the shorter *orfF* (Fig. 7). The protein is present at low levels in LSB-20.1 (which contains ~200 copies of *LD1* on a 54.4-kb circular molecule) and LSB-51.2 (which contains ~10 copies on a 450-kb minichromosome) but is present at much higher levels in LSB-51.1 (with duplicated *FG*x sequence in the *rRNA* locus). The antiserum also cross-reacts with another protein of ~50 kDa in all isolates. Thus, the *FG*x duplication into the *rRNA* locus and transcription with Pol I appears to result in greatly increased gene expression for *orfF* at least.

DISCUSSION

The *LD1* sequence is a 27.2-kb multigene locus occurring within an ~2.2-Mb chromosome in all *Leishmania* isolates examined (34–36, 52, 53). This sequence is amplified, either as small linear chromosomes (200 to 450 kb) or as 54.5-kb circular molecules, in approximately 15% of all *Leishmania* strains tested. Amplification of *LD1* genes is normally accompanied by a corresponding increase in transcript abundance (34, 36). We describe here a novel form of *LD1* amplification: duplication of a single copy of *orfF* and *orfG* into the *rRNA* locus in a 1.2-Mb chromosome, apparently by gene conversion, which results in a substantial increase in their transcript abundance due to transcription by Pol I from the upstream *rRNA* promoter. The increased transcript abundance results in a substantial increase in expression of at least the *orfF* gene product. Thus, overexpression of *orfF* and *orfG* gene products can be accomplished without substantial amplification of the genes, possibly providing a selective advantage to the parasites. The *orfG* gene product has homology to that from *T. brucei ESAG10* (18), and both predicted proteins contain 10 to 12 potential membrane-spanning domains, indicating that they may have a possible transporter function (34). The function of the *orfF* gene product is as yet unknown.

This duplication of *LD1 FG*x sequence in strain LSB-51.1 appears to involve homologous recombination between short (12- or 13-bp) regions within the *LD1* locus on the 2.2-Mb chromosome and the *rRNA* gene locus on the 1.2-Mb chromo-

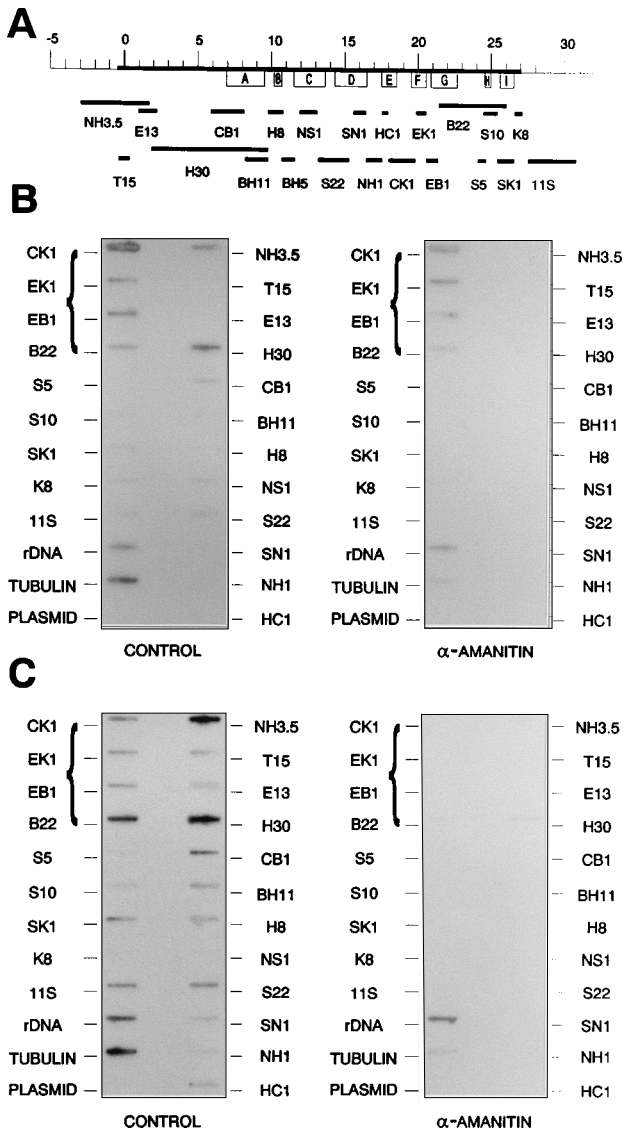

FIG. 6. The duplicated *orfF* and *orfG* in strain LSB-51.1 are transcribed by RNA Pol I. Samples (5 μg) of each plasmid clone representing various segments of the *LD1* locus (A), *T. brucei 18S rRNA* genes (rDNA), *T. brucei* α- and β-tubulin genes (TUBULIN), and pBluescript II SK$^-$ (PLASMID) were blotted onto nylon membranes. Duplicate filters were hybridized with ^{32}P-labeled run-on transcripts from the nuclei of LSB-51.1 (B) or LSB-51.2 (C) labeled in the presence or absence (CONTROL) of 1 mg of α-amanitin per ml. Plasmid designations in the figure are shown minus the initial lowercase p.

some. Homologous recombination in *Leishmania* spp. has been previously demonstrated to play important roles in DNA amplification and development of drug resistance. For example, amplification of P glycoprotein (*ltpgpA*) and short-chain dehydrogenase (*ltdh*) genes within the 30-kb *H* region associated with arsenite and methotrexate resistance in *Leishmania tarentolae* (37, 39) involves intrachromosomal homologous recombination between inverted and direct repeats (198 and 1,241 bp) to form a variety of circular and linear amplicons (21, 38). Generation of a multicopy circular *LD1* molecule (*CD1*) in *L. mexicana* also involves homologous recombination between intrachromosomal direct repeats (28), but in this case the repeats are much shorter (13 bp). The *FG*x duplication in LSB-51.1 also involves homologous recombination between short (12- or 13-bp) repeats, but since the repeats are located on different chromosomes, recombination results in gene conversion rather than the generation of the circular episomes seen with *H*-region and circular *LD1* (*CD1*) amplification. The subsequent (or concomitant) deletion of the ~10-kb segment between *orfG* and sequence x involves homologous recombination between 13-bp direct repeats. Whether this entails intrachromosomal recombination within the duplicated *FG*x region or interchromosomal recombination between the 2.2-Mb *LD1* locus and the duplicated *FG*x region is unknown at this time.

A lower limit in the size of homolog involved in efficient homologous recombination appears to correspond to the respective genome size and is proposed to prevent nonhomologous recombination (17, 49, 54). For example, in mammalian cells, efficient recombination occurs with homologies of >250 bp (26, 41), while in bacterial cells, homologies of 23 to 90 bp (25, 45, 46) are required for efficient recombination. However, low levels of homologous recombination occur in both mammalian and bacterial cells with homologies as small as 16 to 25 bp (2, 41, 46, 54). In addition, sequence deletions have resulted from recombinations involving short (0- to 17-bp) regions of homology (1, 13, 14, 24, 47). The lower limit in homology is thought to be due to the minimal sequence for stable protein-DNA and DNA-DNA interactions (49, 54). The 12- or 13-bp homologies involved in the *FG*x duplication-deletion event(s) reported here are consistent with these reported values, especially when one compares the genome size of *Leishmania* cells ($\sim 5 \times 10^7$) with that of mammalian ($\sim 3 \times 10^9$) and bacterial cells ($\sim 5 \times 10^6$). Thus, a relatively small genome size in *Leishmania* spp. might allow a lower limit in homologies without a concomitant increase in nonhomologous recombinations.

Recombination events have been proposed to occur by two or more pathways with differing efficiencies for short (12- or 13-bp) or long (>100-bp) direct or inverted repeats (54). In *L. donovani*, the association of a 7-bp consensus sequence (tgtG Cat) at or near known junctions of LD1-associated rearrangements suggests that recombinations involving short repeats might be more efficient if associated with a specific recognition site. This argument is strengthened by the presence of one or more copies of this 7-mer within or near the *H*-region repeats involved in recombination events in *L. tarentolae* induced by drug selection (21, 38).

Gene rearrangement and transcription of certain protein-coding genes by Pol I are important mechanisms for regulating transcript abundance in other kinetoplastids (6). The variant

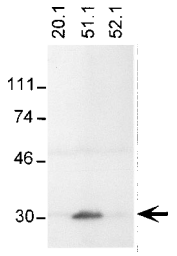

FIG. 7. The *orfF* gene product is overexpressed in strain LSB-51.1. Proteins from 5×10^7 cells of three strains of *L. donovani* (LSB-20.1, LSB-51.1, and LSB-51.2 [shown as 20.1, 51.1, and 52.1, respectively, over the lanes]), were separated by SDS-PAGE, transferred to nitrocellulose filters, and incubated with 1/200 dilution of rabbit antiserum against recombinant *orfF* protein. Bound antibody was visualized by incubation with ^{125}I-labeled protein A and autoradiography. The position of the predicted *orfF* gene product is indicated by the arrow. Molecular mass markers (Gibco-BRL) are shown to the left (in kilodaltons).

surface glycoprotein (*VSG*) gene, *VSG* expression site-associated genes (*ESAG*s), and the procyclic acidic repetitive protein (*PARP*) gene of *T. brucei* are transcribed by Pol I (or a Pol I-like enzyme) (7, 11, 42), and are expressed at high levels. Translation of Pol I-transcribed protein-coding genes is possible in trypanosomatids because of *trans*-splicing of polycistronic pre-mRNAs (33, 48). The 39-nt splice leader (or miniexon) provides the 5′ cap necessary for message transport, stability, and translation (22). Gene conversions involving homologous recombination between sequences within and flanking *VSG* genes are used to regulate *VSG* expression in *T. brucei*, by replacing the *VSG* gene in the active expression site with a previously inactive *VSG* gene (40). Thus, genomic rearrangements which allow transcription of protein-coding genes by Pol I may provide another means for selective regulation of gene expression in *Leishmania* spp. and other members of the Kinetoplastida. Interestingly, a recent report describes the abundant transcription and overexpression of *ltpgpE* in a strain of *L. tropica* selected for methotrexate resistance (16). The increased expression appears to result from a single additional copy of the *ltpgpE* gene, rather than episomal amplification, suggesting that it may involve a gene rearrangement similar to that described in the present study.

ACKNOWLEDGMENTS

We thank Craig Lemley for excellent technical assistance and Lettie Goltry and Dan Hoppe for DNA sequencing.

This work was supported by NIH grant AI17375 from the National Institutes of Health to K.S., who is a Burroughs Wellcome Scholar in Molecular Parasitology; NIH postdoctoral fellowship AI08815 to M.J.L.; a NATO fellowship to G.M.; and a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada to T.D.

REFERENCES

1. **Albertini, A., M. Hofer, M. P. Calos, and J. H. Miller.** 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. Cell **29:**319–328.
2. **Ayares, D., L. Chekuri, K.-Y. Song, and R. Kucherlapati.** 1986. Sequence homology requirements for intermolecular recombination in mammalian cells. Proc. Natl. Acad. Sci. USA **83:**5199–5203.
3. **Beverley, S. M.** 1991. Gene amplification in *Leishmania*. Annu. Rev. Microbiol. **45:**417–444.
4. **Beverley, S. M., J. A. Coderre, D. V. Santi, and R. T. Schimke.** 1984. Unstable DNA amplifications in methotrexate-resistant *Leishmania* consist of extrachromosomal circles which relocalize during stabilization. Cell **38:**431–439.
5. **Beverley, S. M., T. E. Ellenberger, and J. S. Cordingley.** 1986. Primary structure of the gene encoding the bifunctional dihydrofolate reductase-thymidylate synthase of *Leishmania major*. Proc. Natl. Acad. Sci. USA **83:**2584–2588.
6. **Borst, P., and D. R. Greaves.** 1987. Programmed gene rearrangements altering gene expression. Science **235:**658–667.
7. **Brown, S. D., J. Huang, and L. H. T. van der Ploeg.** 1992. The promoter for the procyclic acidic repetitive protein (PARP) genes of *Trypanosoma brucei* shares features with RNA polymerase I promoters. Mol. Cell. Biol. **12:**2644–2652.
8. **Cabot, E. L., and A. T. Beckenbach.** 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput. Appl. Biosci. **5:**233–234.
9. **Carrazana, E., K. B. Pasieka, and J. A. Majzoub.** 1988. The vasopressin mRNA poly(A) tract is unusually long and increases during stimulation of vasopressin gene expression in vivo. Mol. Cell. Biol. **8:**2267–2274.
10. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
11. **Chung, H.-M., M. G.-S. Lee, and L. H. T. van der Ploeg.** 1992. RNA polymerase I-mediated protein-coding gene expression in *Trypanosoma brucei*. Parasitol. Today **8:**414–418.
12. **Devereux, H., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
13. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human β-globin gene family. Cell **21:**653–668.
14. **Farabaugh, P. J., U. Schmeissner, M. Hofer, and J. H. Miller.** 1978. Genetic studies of the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **126:**847–863.
15. **Fourney, R. M., J. Miyakoshi, R. S. Day III, and M. C. Paterson.** 1988. Northern blotting: efficient RNA staining and transfer. Focus **10:**5–7.
16. **Gamarro, F., M. J. Chiquero, M. V. Amador, D. Légaré, M. Ouellette, and S. Castanys.** 1994. *P*-glycoprotein overexpression in methotrexate-resistant *Leishmania tropica*. Biochem. Pharmacol. **47:**1939–1947.
17. **Gonda, D. K., and C. M. Radding.** 1983. By searching progressively RecA protein pairs DNA molecules that share a limited stretch of homology. Cell **34:**647–654.
18. **Gottesdiener, K. M.** 1994. A new VSG expression site-associated gene (ESAG) in the promoter region of *Trypanosoma brucei* encodes a protein with 10 potential transmembrane domains. Mol. Biochem. Parasitol. **63:**143–151.
19. **Greenberg, M. E., and T. P. Bender.** 1990. Identification of newly transcribed RNA, p. 4.10.1–4.10.9. *In* F. W. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, K. Struhl, L. M. Albright, D. M. Coen, and A. Varki (ed.), Current protocols in molecular biology. Current Protocols, New York, N.Y.
20. **Grondal, E. J. M., R. Evers, and A. W. C. A. Cornelissen.** 1990. Identification and sequence analysis of the ribosomal DNA promoter region of *Crithidia fasciculata*. Nucleic Acids Res. **18:**1333–1338.
21. **Grondin, K., B. Papadopoulou, and M. Ouellette.** 1993. Homologous recombination between direct repeat sequences yields p-glycoprotein containing amplicons in assenite resistant *Leishmania*. Nucleic Acids Res. **21:**1895–1901.
22. **Hamm, J., and I. W. Mattaj.** 1990. Monomethylated cap structures facilitate RNA export from the nucleus. Cell **63:**109–118.
23. **Hasan, G., M. J. Turner, and J. S. Cordingley.** 1984. Ribosomal RNA genes of *Trypanosoma brucei*: mapping the regions specifying the six small ribosomal RNAs. Gene **27:**75–86.
24. **Henthorn, P. S., D. L. Mager, T. H. J. Huisman, and O. Smithies.** 1986. A gene deletion ending within a complex array of repeated sequences 3′ to the human β-globin gene cluster. Proc. Natl. Acad. Sci. USA **83:**5194–5198.
25. **Khasanov, F. K., D. J. Zvingila, A. A. Zainullin, A. A. Prozorov, and V. I. Bashkirov.** 1992. Homologous recombination between plasmid and chromosomal DNA in *Bacillus subtilis* requires approximately 70 bp of homology. Mol. Gen. Genet. **234:**494–497.
26. **Liskay, R. M., A. Letsou, and J. L. Stachelek.** 1987. Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. Genetics **115:**161–167.
27. **Liu, J., N. Gajendran, D. Muthui, S. Muyldermans, J. Dujardin, S. De Doncker, D. Jacquet, D. Le Ray, F. Mathieu-Duade, and R. Hamers.** 1991. Chromosome rearrangement in *Leishmania mexicana* M379. Mol. Biochem. Parasitol. **46:**53–60.
28. **Liu, J., G. Salinas, N. Gajendran, D. Muthui, S. Muyldermans, and R. Hamers.** 1992. DNA recombination associated with short direct repeats in *Leishmania mexicana* M379. Mol. Biochem. Parasitol. **50:**351–354.
29. **Liu, X., and K.-P. Chang.** 1992. The 63-kilobase circular amplicon of tunicamycin-resistant *Leishmania amazonensis* contains a functional *N*-acetylglucosamine-1-phosphate transferase gene that can be used as a dominant selectable marker in transfection. Mol. Cell. Biol. **12:**4112–4122.

29a. **Lodes, M. J., et al.** Unpublished data.

30. **Lodes, M. J., B. L. Smiley, A. W. Stadnyk, J. L. Bennett, P. J. Myler, and K. Stuart.** 1993. Expression of a retroposon-like sequence upstream of the putative *Trypanosoma brucei* variant surface glycoprotein gene expression site promoter. Mol. Cell. Biol. **13:**7036–7044.
31. **Milhausen, M., R. G. Nelson, M. Parsons, G. Newport, K. Stuart, and N. Agabian.** 1983. Molecular characterization of initial variants from the IsTat 1 serodeme of *Trypanosoma brucei*. Mol. Biochem. Parasitol. **9:**241–254.
32. **Murphy, N. B., A. Pays, P. Tebabi, H. Coquelet, M. Guyaux, M. Steinert, and E. Pays.** 1987. *Trypanosoma brucei* repeated element with unusual structural and transcriptional properties. J. Mol. Biol. **195:**855–871.
33. **Murphy, W. J., K. P. Watkins, and N. Agabian.** 1986. Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: evidence for *trans* splicing. Cell **47:**517–525.
34. **Myler, P. J., M. J. Lodes, G. Merlin, T. deVos, and K. D. Stuart.** 1994. An amplified DNA element in *Leishmania* encodes potential integral membrane and nucleotide binding proteins. Mol. Biochem. Parasitol. **66:**11–20.
35. **Myler, P. J., C. A. Tripp, L. Thomas, G. M. Venkataraman, G. Merlin, and K. D. Stuart.** 1993. The LD1 amplified element from *Leishmania infantum* encodes a homolog of ribosomal protein L37. Mol. Biochem. Parasitol. **62:**147–152.
36. **Myler, P. J., G. M. Venkataraman, M. J. Lodes, and K. D. Stuart.** 1994. A frequently amplified region in *Leishmania* contains a gene that is conserved in prokaryotes and eukaryotes. Gene **148:**187–193.
37. **Ouellette, M., F. Fase-Fowler, and P. Borst.** 1990. The amplified H circle of methotrexate-resistant *Leishmania tarentolae* contains a novel P-glycoprotein gene. EMBO J. **9:**1027–1033.
38. **Ouellette, M., E. Hettema, D. Wüst, F. Fase-Fowler, and P. Borst.** 1991.

Direct and inverted DNA repeats associated with P-glycoprotein gene amplification in drug resistant *Leishmania*. EMBO J. **10:**1009–1016.
39. **Papadopoulou, B., G. Roy, and M. Ouellette.** 1992. A novel antifolate resistance gene on the amplified H circle of *Leishmania*. EMBO J. **11:**3601–3608.
40. **Pays, E.** 1985. Gene conversion in trypanosome antigenic variation. Prog. Nucleic Acid Res. Mol. Biol. **32:**1–26.
41. **Rubnitz, J., and S. Subramani.** 1984. The minimum amount of homology required for homologous recombination in mammalian cells. Mol. Cell. Biol. **4:**2253–2258.
42. **Rudenko, G., M. G.-S. Lee, and L. H. T. van der Ploeg.** 1992. The PARP and VSG genes of *Trypanosoma brucei* do not resemble RNA polymerase II transcription units in sensitivity to Sarkosyl in nuclear run-on assays. Nucleic Acids Res. **20:**303–306.
43. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
44. **Scholler, J. K., S. G. Reed, and K. Stuart.** 1986. Molecular karyotype of species and subspecies of *Leishmania*. Mol. Biochem. Parasitol. **20:**279–293.
45. **Shen, P., and H. V. Huang.** 1986. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. Genetics **112:**441–457.
46. **Singer, B. S., L. Gold, P. Gauss, and D. H. Doherty.** 1982. Determination of the amount of homology required for recombination in bacteriophage T4. Cell **31:**25–33.
47. **Stringer, J. R.** 1982. DNA sequence homology and chromosomal deletion at a site of SV40 DNA integration. Nature (London) **296:**363–366.
48. **Sutton, R. E., and J. C. Boothroyd.** 1986. Evidence for *trans* splicing in trypanosomes. Cell **47:**527–535.
49. **Thomas, C. A., Jr.** 1966. Recombination of DNA molecules. Prog. Nucleic Acid Res. Mol. Biol. **5:**315–337.
50. **Thomashow, L. S., M. Milhausen, W. J. Rutter, and N. Agabian.** 1983. Tubulin genes are tandemly linked and clustered in the genome of Trypanosoma brucei. Cell **32:**35–43.
51. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76:**4350–4354.
52. **Tripp, C. A., P. J. Myler, and K. Stuart.** 1991. A DNA sequence (LD1) which occurs in several genomic organizations in *Leishmania*. Mol. Biochem. Parasitol. **47:**151–160.
53. **Tripp, C. A., W. A. Wisdom, P. J. Myler, and K. Stuart.** 1992. A multicopy, extrachromosomal DNA in *Leishmania infantum* contains two inverted repeats of the 27.5 kb LD1 sequence and encodes numerous transcripts. Mol. Biochem. Parasitol. **55:**39–50.
54. **Watt, V. M., C. J. Ingles, M. S. Urdea, and W. J. Rutter.** 1985. Homology requirements for recombination in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **82:**4768–4772.
55. **White, T. C., F. Fase-Fowler, H. van Luenen, J. Calafat, and P. Borst.** 1988. The H circles of *Leishmania tarentolae* are a unique amplifiable system of oligomeric DNAs associated with drug resistance. J. Biol. Chem. **263:**16977–16983.