Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 8596–8600, September 1992
Biochemistry

Molecular organization of *Leishmania* RNA virus 1

(RNA-dependent RNA polymerase/*gag–pol* fusion protein/translation frameshift)

KENNETH D. STUART*, REITHA WEEKS†, LYS GUILBRIDE, AND PETER J. MYLER

Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, WA 98109-1651

Communicated by Seymour J. Klebanoff, April 27, 1992

ABSTRACT **The complete 5284-nucleotide sequence of the double-stranded RNA genome of *Leishmania* RNA virus 1 (LRV1) was determined and contains three open reading frames (ORFs) on the plus (+) (mRNA) strand. The predicted amino acid sequence of ORF3 has motifs characteristic of viral RNA-dependent RNA polymerases. ORF2, which may encode the major viral coat protein, overlaps ORF3 by 71 nucleotides, suggesting a +1 translational frameshift to produce a *gag–pol* type of fusion protein. Two alternative models for the frameshift are presented. The 5′ splice leader sequence of kinetoplastid mRNAs is not in LRV1 RNA. This suggests that the 450-base region at the 5′ end of the LRV1 (+)-strand, which contains ORF1 and is highly conserved among viral strains, does not encode protein but has a role in initiation of translation and/or RNA stability. The similarity of LRV1 genomic organization, replication cycle, and RNA-dependent RNA polymerase sequence to those of the yeast virus ScV L-A suggests a common ancestral origin. The possibility that LRV1 affects pathogenesis in leishmaniasis is intriguing.**

A small, but growing, number of viruses have been studied in parasites. There are DNA viruses in *Entamoeba histolytica* and double-stranded (ds) RNA viruses in *Babesia*, *Eimeria*, *Giardia*, *Leishmania*, and *Trichomonas* (1). The *Giardia* and *Eimeria stiedae* viruses cross-hybridize (2). *Entamoeba* and *Giardia* produce infective viral particles, and *Giardia* has been transfected with single-stranded (ss) viral RNA (1). *Trichomonas*, *Eimeria*, and *Leishmania* do not appear to produce infective viral particles (1, 3).

Leishmania RNA virus 1 (LRV1) was discovered in *Leishmania guyanensis* (4, 5). All 12 viral isolates found to date are in *L. braziliensis* and *L. guyanensis* from the Amazon basin (6, 7), although several (5/17) isolates from this region do not contain virus. Six LRV1 cDNA clones, representing about a third of the viral genome, hybridize to the other 11 viral RNAs (7), which thus appear to be strains of the same virus.

Leishmania infected with LRV1 contain dsRNA and substantially less (+)-strand RNA (3). We determined the entire nucleotide sequence of LRV1 from cDNAs and found two large and one small open reading frame (ORF) on the (+)-strand.‡ ORF3 encodes RNA-dependent RNA polymerase (RDRP) and overlaps ORF2, which may encode the viral coat protein. This implies that a +1 translational frameshift produces a *gag–pol* type of fusion protein.

MATERIALS AND METHODS

Culture of Parasites and Isolation of RNA. Promastigotes from *L. guyanensis* MHOM/SR/80/CUMC1 clone 1A were grown to late logarithmic or stationary phase (10^8 cells per ml) before harvest (5). Total RNA was prepared from *Leishmania* by urea/phenol or guanidinium thiocyanate–phenol/chloroform extraction (3, 7). The 5.3-kilobase (kb) LRV1 dsRNA was purified by electrophoresis in 0.6% agarose/TBE (0.09 M Tris borate/2 mM EDTA, pH 8.0) gels, electroelution, and ethanol precipitation.

LRV1 Cloning. LRV1 cDNA clones were prepared from gel-purified RNA (5) or by PCR amplification. Their location within the LRV1 genome is shown in Fig. 1. First-strand cDNA was synthesized using mixed hexamer primers (LP series), oligo(dT) priming of *Escherichia coli* poly(A) polymerase-treated LRV1 RNA (LJ series); and oligonucleotides 87-19 and 87-20 (see Fig. 2) (LW or WAL series). Second strand was synthesized by the RNase H method (9), repaired, methylated, linkered, and ligated into *Bam*HI-digested pBS$^+$ (LP series), *Eco*RI-digested λZAP (LJ series) or *Eco*RI-digested λZAPII (LW and WAL series) (10). Clones of 5′ ends were obtained by PCR amplification of C-tailed (terminal deoxynucleotidyltransferase) first-strand cDNA synthesized using primers 90-122 and 87-19 for the (+)- and (−)-strand, respectively. Amplification with Bam-dG_{10} (CCGGATCCGGGGGGGGGG) and 90-145 or 89-195, respectively, used Replinase (DuPont/NEN) for 30 cycles of denaturation at 94°C for 1 min, annealing at 45°C for 1 min, and extension at 72°C for 2 min. The products were digested with *Bam*HI and *Eco*RI (1P145 series), *Bam*HI and *Nhe* I (1PN4 series), or *Bam*HI and *Hin*dIII (1P195 series) and ligated into pBluescript II SK(−) (Stratagene) digested with the appropriate enzymes. For cloning the 3′ ends, gel-purified dsRNA was C-tailed using poly(A) polymerase (BRL), cDNA was synthesized using Bam-dG_{10}, and PCR amplification was carried out using Bam-dG_{10} and 89-195 [(+)-strand] or 90-122 followed by reamplification with 90-145 and Bam-dG_{10} [(−)-strand]. The products were digested with *Bam*HI and *Hin*dIII (2P195 series) or *Bam*HI and *Nhe* I (2PN4 series) and cloned as above. Primers 90-122 and 90-146 were used for first-strand cDNA synthesis and PCR amplification of gel-purified LRV1 dsRNA for some internal LRV1 clones. Products were digested with *Eco*RI and ligated into *Eco*RI-digested pBluescript II SK(−) (PCR146-1 and P2R series).

Sequence Analysis. Plasmid and PCR product DNA was sequenced by the dideoxy chain-termination method using Sequenase (United States Biochemical) (11) and sequence analysis was carried out using DNASTAR and PCFOLD (12) and CLUSTAL (13) software. Homology searches of the Swiss-Prot (Release 19.0) and GenPept (Release 64.3) data bases used the FASTA program (14).

RESULTS

Complete Nucleotide Sequence of LRV1. The 5284-nucleotide (nt) consensus sequence of the LRV1 genome

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: LRV1, *Leishmania* RNA virus 1; ss, single-stranded; ds, double-stranded; ORF, open reading frame; RDRP, RNA-dependent RNA polymerase; SL, splice leader; nt, nucleotide(s).

*To whom reprint requests should be addressed.
†Present address: Bristol-Myers Squibb Pharmaceutical Research Institute, 3005 First Avenue, Seattle, WA 98121.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M92355).

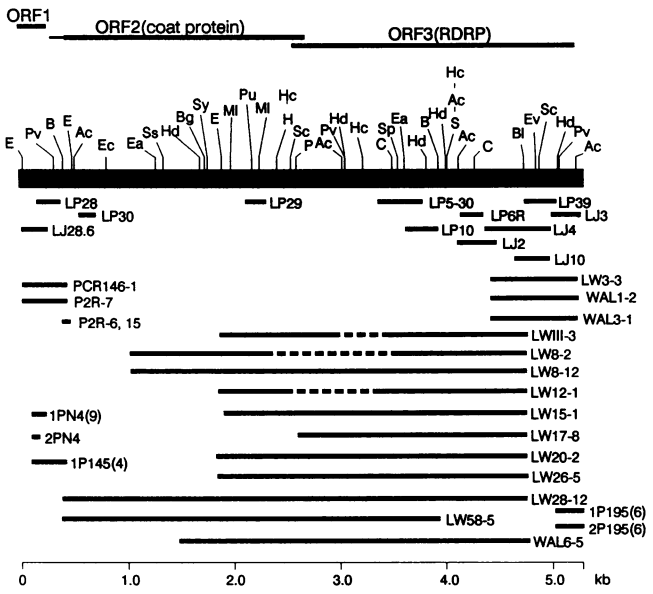

FIG. 1. Genomic organization of LRV1. The location of LRV1 cDNA and PCR clones are shown below the restriction map derived from these clones (Ac, *Acc* I; B, *Bam*HI; Bg, *Bgl* I; Bl, *Bgl* II; C, *Cla* I; E, *Eco*RI; Ea, *Eag* I; Ec, *Eco*RV; H, *Hpa* I; Hc, *Hin*cII; Hd, *Hin*dIII; MI, *Mlu* I; P, *Pst* I; Pv, *Pvu* I; Pu, *Pvu* II; Sc, *Sca* I; Sp, *Sph* I; Ss, *Sst* I; Sy, *Sty* I). Dashed lines indicate deletions within some cDNA clones. Thick lines indicate the position and sizes of ORFs beginning with AUG initiation codons; thin lines at the 5′ ends of the ORFs indicate the distance to the next upstream termination codon.

(Fig. 2) was determined from a series of overlapping independent cDNA clones (Fig. 1) of LRV1 RNA from *L. guyanensis* MHOM/SR/80/CUMC1-1A. Of 19 differences between cDNA clones, 12 were between clones from libraries prepared at three times over a 2-year period, during which substantial parasite cultivation occurred. Thus, the differences may primarily represent viral replication errors. Nine of these differences involve U → C or A → G transitions reminiscent of unwinding–modifying activity (17), but no such activity was detected in *Leishmania* extracts (B. Bass, personal communication). The remaining differences may reflect sequence diversity in the original virus population or may have arisen during cDNA cloning.

There appears to be some heterogeneity in the length of each viral RNA strand (Fig. 3), although differences due to cloning artifacts are difficult to exclude. Interestingly, the 3′ termini of each strand tend to be longer than the 5′ ends of the opposite strand, suggesting that LRV1 dsRNA may have staggered ends. The 3′ end of the (+)-strand also contains several tandem direct repeats of the trinucleotide sequence TCA. A single copy of this sequence also occurs at the 3′ end of the (−)-strand, suggesting a possible role in replication of LRV1 RNA. The locations of the termini were confirmed by primer extension analysis and by measuring the size of end-labeled fragments released by digestion of LRV1 RNA with RNase H following hybridization with oligonucleotides situated near the 5′ and 3′ termini of both strands (unpublished data).

LRV1 Encodes an RDRP. Analysis of the LRV1 consensus sequence reveals two large and one small ORF, each in a different reading frame of the same RNA strand (Fig. 2). The complementary strand contains numerous termination codons in all three reading frames, with no ORFs of >450 nt. The coding strand corresponds to the LRV1 ssRNA present in infected cells, and thus this ssRNA is the (+)-strand that can be translated into protein. The predicted amino acid sequence from ORF3 contains the eight motifs conserved in the C-terminal domain of RDRPs from other (+)-strand and dsRNA viruses (16), suggesting that this ORF encodes the RDRP activity detected in virus-infected parasites (15, 18). CLUSTAL (13) alignment of several viral RDRP sequences (data not shown) indicates that LRV1 ORF3 is most closely related to the RDRP from ScV L-A (or L1), a dsRNA virus of yeast (19, 20). The homology between the LRV1 and ScV RDRPs extends beyond the region conserved among other viruses (Fig. 2), with 23% identical amino acids, 27% conservative replacements, and 15% neutral replacements for a total of 64% overall similarity. The predicted size (beginning with AUG, nt 2622) (98 kDa) and pI (8.2) of the ORF3 product are similar to those of the yeast ScV L-A RDRP (99 kDa and 8.6, respectively), and the two predicted proteins have similar hydropathic profiles (unpublished data). The termination codon of ORF3 at nt 5233 is 48 nt upstream of the terminus of the LRV1 RNA.

In other RNA viruses, including ScV L-A, the viral coat protein gene is often upstream of the RDRP gene (21, 22). ORF2 is upstream of the RDRP gene in LRV1 and predicts a 741-amino acid (82-kDa) protein with pI 6.5 (beginning with AUG, nt 448). Although coat protein sequences are not well conserved among RNA viruses, the predicted size and pI are similar to those of the ScV L-A coat protein (76 kDa and 5.6). Data base searches (see *Materials and Methods*) showed no significant sequence similarity to any protein in the data bases. (See *Note Added in Proof.*)

Potential +1 Translational Frameshift. LRV1 ssRNA (mRNA) is similar, if not identical, in length to the dsRNA (3), and other viral genome-sized mRNAs often use translational frameshifting (23) to synthesize a *gag–pol* type of fusion protein that appears to assemble in viral particles and provide RDRP activity and RNA binding sites (21). The 71-nt (nt 2600–2670) overlap between LRV1 ORFs 2 and 3 (Fig. 2) suggests a +1 translational frameshift to produce an analogous fusion protein of ≈176 kDa with pI ≈ 7.6. The sequence CCCGAAC at nt 2620 may allow "slippage" after tRNA decoding, resulting in a +1 frameshift analogous to that seen in yeast Ty elements (24). In this model, mispairing of the "slippery" $tRNA^{Pro}_{CGG}$ at the CCC codon (nt 2620) and subsequent stalling of translation by the predicted stem/loop structure (nt 2625–2643) would permit the $tRNA^{Pro}_{CGG}$ to slip from CCC to the CCG codon, resulting in a +1 frameshift. A low abundance of the $tRNA^{Glu}_{UUC}$ suggested by the low rate of usage of the GAA codon in *Leishmania* (S. Beverley, personal communication) may also facilitate translational stalling.

Alternatively, a larger potential secondary structure might produce the +1 frameshift by the ribosomal bypass or "long jump" mechanism described for bacteriophage T4 gene *60* (25). In this model, the GAU codon at nt 2611 just upstream of the stem/loop structure (nt 2618–2653; Fig. 4) predicted by PCFOLD ($\Delta G = -20.3$ kcal/mol) would be the "take-off" site, while another GAU codon, at nt 2672, would be the "landing" site. However, the sequence has the potential to form a pseudoknot structure (Fig. 4) by base pairing between the CCAC at nt 2632 and the GUGG at nt 2660 (to form stem III), which may allow further base pairing between the CUGCG at nt 2664 with the UGCAG at nt 2613 (stem IV). The asymmetric mismatch between stems I and II may act as a hinge, and stacking of A residues within the hinge and loop, along with formation of two G·U base pairs (V) may stabilize the pseudoknot. Formation of this pseudoknot would bring the two GAU codons into close physical proximity and could facilitate binding of the $tRNA^{Asp}_{AUC}$ at the downstream GAU in the +1 reading frame rather than at the upstream GAU.

5′ End of LRV1 (+)-Strand. The short ORF near the 5′ end of the LRV1 (+)-strand (ORF1, beginning with AUG, nt 15) predicts a 71-amino acid (7.8-kDa, pI 4.9) polypeptide with no significant similarity to any protein in the data bases. However, PCFOLD analysis (12) predicts considerable secondary

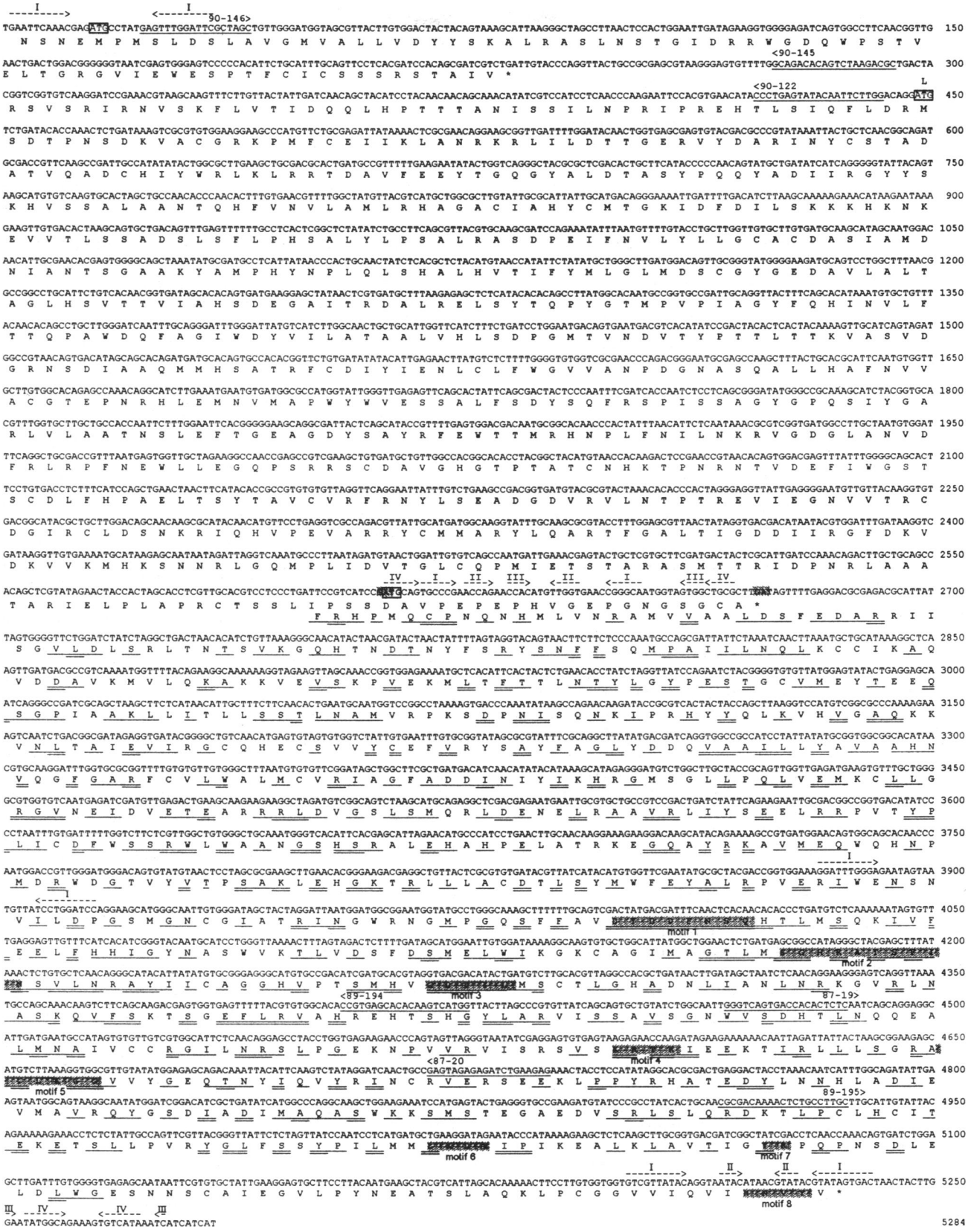

FIG. 2. Nucleotide and deduced amino acid sequence of LRV1. The 5284-nt LRV1 consensus sequence and the amino acid sequences of the three ORFs from the (+)-strand are shown. Potential AUG initiation codons are boxed, and codons potentially involved in translational frameshift between ORF2 and ORF3 are shaded. Amino acid identities with RDRP from ScV L-A (15) are indicated by double underlining, conservative replacements by single underlining, and the eight motifs characteristic of viral RDRPs (16) are shaded. Oligonucleotides used in the synthesis of cDNA clones are shown by underlining and their orientation indicated by arrowheads; 90-122, 90-145, and 90-146 contain an *Eco*RI linker (CGGAATTC) at their 5′ end. Complementary sequences with potential to form the secondary structures discussed are indicated by arrows and roman numerals.

```
   1        10        20        30        40        50          5'(+)    3'(-)
   |        |         |         |         |         |
5' TGAATTCAAACGAGATGCCTATGAGTTTGGATTCGCTAGCTGTTGGGATG----        2        4
   CGAATTCAAACGAGATGCCTATGAGTTTGGATTCGCTAGCTGTTGGGATG----        0        1*
    GAATTCAAACGAGATGCCTATGAGTTTGGATTCGCTAGCTGTTGGGATG----        9        2
          AAACGAGATGCCTATGAGTTTGGATTCGCTAGCTGTTGGGATG----        1        0
                                              TTGGGATG----        1        0

total                                                          13        7


    5240      5250      5260      5270      5280                3'(+)    5'(-)
    |         |         |         |         |
----CTAACTACTTGGAATATGGCAGAAAGTGTCATAAATCATCATCAT 3'            2        0
----CTAACTACTTGGAATATGGCAGAAAGTGTCATAAATCATCA                   0        1
----CTAACTACTTGGAATATGGCAGAAAGTGTCATAAATCATCT                   2        0
----CTAACTACTTGGAATATGGCAGAAAGTGTCATAAATCAT                     1        5
----CTAACTACTTGGAATATGGCAGAAAGTGTCATAAATCA                      1        0

total                                                          6        6
```

FIG. 3. The 5′ and 3′ termini of LRV1 are heterogeneous. The number and nucleotide sequence of PCR and conventional (indicated by the asterisk) clones containing terminal sequence from ds LRV1 are shown. The numbers above the sequence indicate the nucleotide position.

structure in this region of LRV1 (unpublished data), and sequence conservation (discussed below) suggests a role other than protein coding.

The labeling efficiency of the 5′ ends of LRV1 dsRNA with T4 polynucleotide kinase was increased by pretreatment with calf intestinal alkaline phosphatase (unpublished data), confirming results obtained by others (6). Thus, both 5′ ends of isolated RNA appear to contain 5′ phosphates (probably 5′ monophosphates), rather than the typical eukaryotic cap structures found on nuclearly encoded *Leishmania* mRNAs. The 5′-terminal sequences of both LRV1 strands, cloned from viral dsRNA, have no homology to the 39-nt splice leader (SL) sequence present on the 5′ terminus of all kinetoplastid mRNAs (8). When cDNA synthesized from total cellular RNA from LRV1-infected *L. guyanensis* using oligonucleotide 90-122 was amplified by PCR using 90-122 primer and oligonucleotide X-SL24 (TATCTAGAACAGTTTCTGTACTATATTG), which matches the SL sequence, no discrete product was observed (unpublished data). Parallel amplifications using oligonucleotides 90-122 and 90-146 or X-SL24 with oligonucleotides specific for other cellular mRNAs produced fragments of the appropriate sizes. Thus, the SL sequence is absent from the 420 nt at the 5′ end of LRV1 RNA. The SL sequence might be added to LRV1 ssRNA 3′ to nt 422–442 and 5′ to the ORF2 initiation codon (nt 448), but this is unlikely given the full genome length of LRV1 ssRNA (3).

Other Potential Secondary Structures. PCFOLD analysis (12) of the LRV1 (+)-strand using a search window of 100 nt and maximum loop size of 15 nt predicts numerous secondary structures. A 10-base-pair (bp) stem/13-nt loop structure that contains the AUG initiation codon of ORF1 is predicted at the 5′ terminus (nt 2–34; Fig. 2). At the 3′ terminus, a large stem/loop structure, containing the termination codon of ORF3, is predicted at nt 5201–5240, while a smaller stem/loop structure is predicted at nt 5251–5276 (Fig. 2). Only one other large (>9-bp) stem/loop structure (nt 3882–3915; Fig. 2) is predicted, although a number of smaller structures are predicted. While similar structures ≈400 nt from the 3′ terminus of yeast ScV L-A, X, and M_1 RNAs are important in binding of the ssRNA by the RDRP fusion protein prior to viral assembly (26), the physiological significance of the predicted structures in LRV1 remains to be explored.

DISCUSSION

We determined the entire 5284-nt sequence of LRV1 and showed that one strand contains three ORFs, each in a different frame. The other strand contains no ORFs longer than 450 nt. The strand with the three ORFs is present in infected parasites as ssRNA (3), suggesting that this is the viral (+)-strand, which is translated. The 3′ ORF3 predicts a protein with several motifs that are conserved among viral RDRPs (16), suggesting that it is responsible for the RDRP activity detected in virus-infected parasites (15, 18). The putative LRV1 RDRP (ORF3) has greatest similarity (64%) to that of the yeast dsRNA virus ScV L-A, extending beyond the conserved motifs. Since LRV1 ORF2 is upstream of, and overlaps, the RDRP (ORF3), it probably encodes the major viral coat protein, as in other RNA viruses (21). The predicted molecular mass and pI are similar to those of the ScV L-A coat protein (81.7 kDa vs. 76.0 kDa, and 6.52 vs. 5.62). No homology was detected to the ScV L-A major coat protein or any protein in the GenPept and Swiss-Prot data bases, but sequences are not well conserved among coat proteins of RNA viruses.

The 5′ ORF1 of the LRV1 (+)-strand could encode a small acidic protein. ScV L-A contains no equivalent ORF, although the killer toxins encoded by dsRNA in yeast and fungi (27) have a size similar to the potential ORF1 product and small (2.4- to 12-kDa) genome-linked viral proteins (VPg) have been found covalently linked to 5′-terminal uridines in some viruses (28). It is unlikely that ORF1 encodes a VPg, since they are generally basic with little helical structure, unlike the protein predicted by ORF1. The possibility of

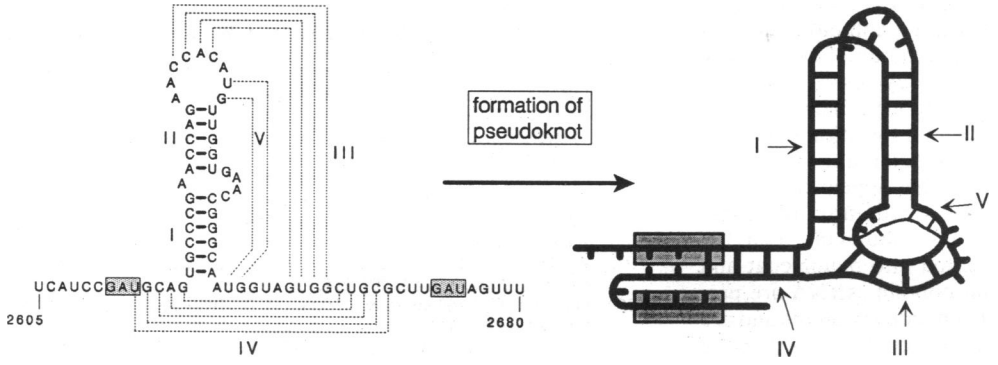

FIG. 4. LRV1 RNA may contain a pseudoknot structure. At left is shown the potential secondary structure in LRV1 RNA, nt 2605–2680. Dotted lines indicate base pairs formed during pseudoknot formation, which is diagrammed at right. The proposed GAU codon ribosomal "take-off" and "landing" sites are boxed.

translational frameshifting producing an ORF1/ORF2 fusion protein is unlikely because the region between the ORFs contains stop codons in all reading frames. The lack of a typical eukaryotic cap structure at the 5′ end of LRV1 (+)-strand RNA and the absence of the SL sequence, both of which are normally required for efficient ribosome binding and translations in trypanosomatids, suggest that sequences 5′ to ORF2 may serve to initiate translation by ribosome binding at internal sites as in other viral systems (29). Picornaviruses initiate translation of their uncapped mRNAs by ribosome binding to an internal site (29, 30). This mechanism requires a 450-nt sequence at the 5′ terminus of the mRNA that appears to have a high degree of secondary structure. Thus, it is possible that the corresponding region of LRV1, which includes ORF1, serves a similar function. These sequences may also have a role in stabilization of viral RNA. The finding that this region contains the most highly conserved sequence among different LRV1 viral strains (7) supports a role at the nucleic acid, rather than amino acid, level. Resolution of these questions requires further study.

By analogy to other viral systems, particularly ScV L-A, it seems likely that translational frameshifting may produce a *gag–pol* type of fusion protein from ORFs 2 and 3. Production of the fusion protein at a low molar ratio to the major coat protein would provide for assembly of viral particles that contain one or a few proteins with RDRP activity and perhaps with RNA binding sites (21). Two alternative models for translational frameshifting between ORF2 and ORF3 are possible. The tRNA slippage model is analogous to that seen in yeast Ty elements (24). The abundance of "slippery" $tRNA^{Pro}_{CGG}$ may control the frequency with which the translational frameshift occurs. The ribosomal bypass model may be more likely, due to the presence of a large and stable predicted stem/loop structure that could provide the predicted translation stall (23, 25). The possibility that pseudoknot formation (Fig. 4) may position the downstream +1 GAU codon so that it rather than the upstream GAU is decoded is intriguing. This frameshift may be a "short hop," rather than the "long jump" proposed for bacteriophage T4 gene *60* (25). Low frequency of pseudoknot formation may result in few fusion proteins relative to coat proteins.

The secondary structure of the LRV1 (+)-strand ssRNA predicted by PCFOLD is complex, with numerous possible stem/loop structures. Stem/loops at the 5′ and 3′ termini may function in initiation of translation and RNA stabilization and replication (31), respectively. In the yeast killer system, stem/loop structures ≈400 nt from the 3′ terminus of ScV L-A, X, and M_1 RNAs were found to be important in binding of the ssRNA by the RDRP fusion protein prior to viral assembly (26).

The lack of detectable sequence homology between LRV1 and the (−)-strand RNA viruses from sandflies (32), the insect host of leishmanial parasites, indicates that LRV1 is unlikely to be derived from viruses of the insect vector. The similarity to the yeast ScV L-A and L-BC RNAs (21, 33), in terms of RDRP amino acid sequence, genome organization, and replication cycle, suggests that LRV1 is an endogenous virus of *Leishmania* that diverged from a common ancestor with ScV. LRV1 and ScV have a number of features in common and appear to have similar life cycles (3, 34). They have similar genomic organizations, although their 5′ termini differ, perhaps reflecting differences in translation initiation. Their unsegmented dsRNA genomes are contained in similar-sized cytoplasmic isometric particles that appear to have no envelope (1). The dsRNA is transcribed into full-length (+)-strands (15, 34, 35) that are less abundant than the dsRNA and are presumably translated into major coat and coat/RDRP proteins. Particles that contain ssRNA are presumably assembled from the virally encoded proteins and are less dense and have lower sedimentation coefficients than the dsRNA particles (3, 35, 36). These ssRNA particles synthesize (conservatively replicate) the (−)-strand, making dsRNA (35), thus completing the cycle. In addition, LRV1 does not appear to produce extracellular infective viral particles (3) but may be transmitted by cell contact and perhaps mating, like ScV (21). In yeast, other dsRNAs, some of which encode toxins, are associated with the viral particles (21). This raises the possibility that similar RNAs may also exist in *Leishmania*, possibly affecting the host–parasite interaction, causing differences in virulence and pathology.

Note Added in Proof. We find homology between the amino acid sequences predicted from LRV1 ORF2 and a cDNA encoding part of the 100-kDa presumptive coat protein of *Eimeria stiedae* virus (ESV) (provided by H. Revets), suggesting that ORF2 encodes LRV1 coat protein.

Contributions by R.W. and L.G. were essentially equivalent. We thank W. Wisdom, P. Tarr, and J. Scholler for various contributions to this work. This investigation received financial support from the United Nations Development Program/World Bank/World Health Organization Special Program for Research and Training in Tropical Diseases and National Institutes of Health Grant AI124771. K.D.S. is a Burroughs Wellcome Scholar in Molecular Parasitology.

1. Wang, A. L. & Wang, C. C. (1991) *Parasitol. Today* **7**, 76–80.
2. Revets, H., Dekegel, D., Deleersnijder, W., de Jonckheere, J., Peeters, J., Leysen, E. & Hamers, R. (1989) *Mol. Biochem. Parasitol.* **36**, 209–215.
3. Weeks, R., Aline, R. F., Jr., Myler, P. J. & Stuart, K. (1992) *J. Virol.* **66**, 1389–1393.
4. Stuart, K., Tarr, P., Aline, R., Smiley, B. & Scholler, J. (1987) in *Molecular Genetics of Protozoa*, ed. Turner, M. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 158–163.
5. Tarr, P. I., Aline, R. F., Jr., Smiley, B. L., Scholler, J., Keithly, J. & Stuart, K. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9572–9575.
6. Widmer, G., Comeau, A. M., Furlong, D. B., Wirth, D. F. & Patterson, J. L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5979–5982.
7. Guilbride, L., Myler, P. J. & Stuart, K. (1992) *Mol. Biochem. Parasitol.* **54**, 101–104.
8. Perry, K. & Agabian, N. (1991) *Experientia* **47**, 118–128.
9. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269.
10. Sambrook, J., Fritsch, T. & Maniatis, T. (1987) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
11. Tabor, S. & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4767–4771.
12. Zuker, M. (1989) *Methods Enzymol.* **180**, 262–288.
13. Higgins, D. G. & Sharp, P. M. (1988) *Gene* **73**, 237–244.
14. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
15. Widmer, G. & Patterson, J. L. (1991) *J. Virol.* **65**, 4211–4215.
16. Bruenn, J. A. (1991) *Nucleic Acids Res.* **19**, 217–226.
17. Bass, B. L., Weintraub, H., Cattaneo, R. & Billeter, M. A. (1989) *Cell* **56**, 331.
18. Widmer, G., Keenan, M. C. & Patterson, J. L. (1990) *J. Virol.* **64**, 3712–3715.
19. Icho, T. & Wickner, R. B. (1989) *J. Biol. Chem.* **264**, 6716–6723.
20. Diamond, M. E., Dowhanick, J. J., Nemeroff, M. E., Pietras, D. F., Tu, C. & Bruenn, J. A. (1989) *J. Virol.* **63**, 3983–3990.
21. Wickner, R. B. (1989) *FASEB J.* **3**, 2257–2265.
22. Strauss, J. H. & Strauss, E. G. (1988) *Annu. Rev. Microbiol.* **42**, 657–683.
23. Atkins, J. F., Weiss, R. B. & Gesteland, R. F. (1990) *Cell* **62**, 413–423.
24. Belcourt, M. F. & Farabaugh, P. J. (1990) *Cell* **62**, 339–352.
25. Weiss, R. B., Huang, W. M. & Dunn, D. M. (1990) *Cell* **62**, 117–126.
26. Fujimura, T., Esteban, R., Esteban, L. M. & Wickner, R. B. (1990) *Cell* **62**, 819–828.
27. Tao, J., Ginsberg, I., Banerjee, N., Held, W., Koltin, Y. & Bruenn, J. A. (1990) *Mol. Cell. Biol.* **10**, 1373–1381.
28. Wimmer, E. (1982) *Cell* **28**, 199–201.
29. Jackson, R. J., Howell, M. T. & Kaminski, A. (1990) *Trends Biochem. Sci.* **15**, 477–483.
30. Evstafieva, A. G., Ugarova, T. Y., Chernov, B. K. & Shatsky, I. N. (1991) *Nucleic Acids Res.* **19**, 665–671.
31. Esteban, R., Fujimura, T. & Wickner, R. B. (1989) *EMBO J.* **8**, 947–954.
32. Bishop, D. H. L. (1986) in *Fundamental Virology*, eds. Fields, B. N. & Knipe, D. M. (Raven, New York), pp. 449–478.
33. Wickner, R. B. (1986) *Annu. Rev. Biochem.* **55**, 373–395.
34. Weeks, R. S., Patterson, J. L., Stuart, K. & Widmer, G. (1991) *Mol. Biochem. Parasitol.* **52**, 207–214.
35. Fujimura, T., Esteban, R. & Wickner, R. B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4433–4437.
36. Fujimura, T. & Wickner, R. B. (1987) *Mol. Cell. Biol.* **7**, 420–426.