Identification and Visualization of the Dimerization Initiation Site of the Prototype Lentivirus, Maedi Visna Virus: A Potential GACG Tetraloop Displays Structural Homology with the α- and γ-Retroviruses†

Tom P. Monie,[△,‡,§] Jane S. Greatorex,[△,‡] Laura Maynard-Smith,[‡] Ben D. C. Hook,[‡] Naomi Bishop,[‡,∥] Lucy P. Beales,[⊥,#] and Andrew M. L. Lever*[,‡]

Department of Medicine, University of Cambridge, Level 5, Addenbrooke's Hospital, Cambridge CB2 2QQ, U.K., and School of Biochemistry and Molecular Biology, University of Leeds, Division of Microbiology Old Medical School, Thoresby Place, Leeds LS2 9JT, U.K.

Received July 12, 2004; Revised Manuscript Received September 17, 2004

ABSTRACT: Dimerization of retroviral genomic RNA is essential for efficient viral replication and is mediated by structural interactions between identical RNA motifs in the viral leader region. We have visualized, by electron microscopy, RNA dimers formed from the leader region of the prototype lentivirus, maedi visna virus. Characterization by in vitro assays of the domains responsible for this interaction has identified a 20 nucleotide sequence that functions as the core dimerization initiation site. This region is predicted to form a GACG tetraloop and therefore differs significantly from the kissing loop palindromes utilized to initiate dimerization in primate lentiviruses. The motif is strongly conserved across the ovine and caprine lentiviruses, implying a critical functional role. Furthermore, the proposed GACG tetraloop exhibits marked structural homology with similar structural motifs present in the leader regions of the α- and γ-retroviruses, and the maedi visna virus dimer linkage region is capable of forming heterodimeric species with the Moloney murine leukemia virus Ψ domain. This may be indicative of commonality of origin of the two viruses or convergent evolution.

Retroviral RNA genomes possess highly structured untranslated RNA leader regions containing a complex array of RNA structural motifs displaying *cis*- and *trans*-regulatory functions. The dimerization initiation site (DIS), responsible for mediating the formation of retroviral genomic RNA dimers, is one such motif. A dimeric retroviral genome is a critical requirement in achieving efficient viral infectivity and replication, with the formation of RNA dimers influencing both viral RNA packaging (*1*−*4*) and reverse transcription (*5*, *6*). The importance of RNA dimerization is emphasized by the significantly impaired replication of the human immunodeficiency virus type 1 (HIV-1) when the DIS is mutated (*5*, *7*).

Retroviral RNA dimers result from noncovalent interactions between identical structural motifs near the viral 5′ termini (*4*, *8*). The accepted mechanism of dimer formation appears to be an initial kissing loop interaction between autocomplementary sequences presented on RNA stem−loops (*9*−*13*) resulting in the formation of immature, or low stability, dimers, which mature into more stable extended duplexes (*14*).

RNA dimerization in the primate lentiviruses, predominantly HIV-1, has been extensively studied, yet little has been published on this process in the nonprimate lentiviruses. However, studies of rapid harvest virions of the prototype lentivirus, maedi visna virus (MVV), identified viral RNA with a Svedberg coefficient of 35 S immediately postbudding, increasing with time to 70 S (*15*). Although weakly interacting dimers formed during RNA encapsidation may have been denatured during purification, these observations are supportive of a progression from monomeric to dimeric RNA concomitant with viral maturation.

We investigated the structure of the MVV leader RNA using a combination of free energy minimization studies, direct visualization by negative stain transmission electron microscopy, and biochemical analysis of dimer formation in truncated transcripts. Inhibition of dimerization by targeted DNA oligonucleotides and site-directed mutagenesis confirmed the location and structure of the MVV DIS, and we noted structural and functional homology of this region with similar regions in distantly related retroviruses.

EXPERIMENTAL PROCEDURES

Plasmid Construction. The transcription plasmid pGVVS-N+S possesses a pGEM11 (Promega) backbone with a

† T.P.M. was supported by a Medical Research Council studentship.

* Corresponding author. Tel: +44-1223-330191. Fax: +44-1223-763401. E-mail: amll1@mole.bio.cam.ac.uk.

△ These two authors contributed equally to this work.

‡ University of Cambridge.

§ Current address: Biophysics Section, Department of Biological Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, U.K.

∥ Current address: 2205 Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, U.K.

⊥ University of Leeds.

Current address: Department of Microbiology and Immunology, University of Texas Medical Branch at Galveston, Galveston, TX 77555-1019.

10.1021/bi048529m CCC: $30.25 © 2005 American Chemical Society
Published on Web 12/07/2004

Table 1: Primers Used To Generate the Transcripts

start site	forward primer	primer sequence (5′ → 3′)
1	MVVT7 1−23	TTGGCGCGCCTAATACGACTCACTATAGGGAAAAGCAGAGTGCTTTGGAG
98	MVVT7 98	TTGGCGCGCCTAATACGACTCACTATAGCTTGCTG GTTATTATCGGGATT
144	MVVT7 144	TAATACGACTCACTATAGGAGCGGATCTCGCAGC
212	MVVT7 212	TAATACGACTCACTATAGGGACCACGGACGCTGC
271	MVVT7 271	TAATACGACTCACTATAGGGAAAGAAAGCTTCGG
303	MVVT7 303	TAATACGACTCACTATAGGTAAGAGAGACACCTA

RNA transcript	reverse primer	primer sequence (5′ → 3′)
MVV1−571	571−562	CCCTACCCTTATTTTACAAG
MVV1−485	485−468*	TCTAGCTAGCCTCACCAC
MVV1−400	400−383*	TTTCTAAGGGGTCGCGAA
MVV1−350	350−330*	TCACTGAAGGGCTCCTCGCCG
MVV1−303	303−281*	CTTACTTCAGGCGTCCCCGAAGC
MVV1−300	300−283	ACTTCAGGCGTCCCCGAA
MVV1−280	280−253*	TTTCTTTCACTTTTTGTCCGTGTCCG
MVV1−260	260−239*	GTCCGTTACTCTCTCCTCGCCG
MVV1−220	220−203*	CGTGGTCCCATAACTCTC
MVV1−159	159−138	CTGCGAGATCCGCTCCGGTGTT
MVV1−97	97−77	TCAGATATCAGCCAACTCCTT
N/A	382−348*	CCCTCCTCCTGTGCCCAAGCAACACTTTCTCTTTC
N/A	329−304*	ACTTCCCCAGTAGGTGTCTCTCTTAC
N/A	239−221*	TCCTCACAGGAGCAGCGTC

significantly deleted multicloning site. The first 480 nucleotides of the MVV_{KV1772} leader RNA [numbering according to Andresson et al. (*16*)] were inserted into the remaining *Nhe*I restriction site. A second restriction site was reintroduced by partial digestion with *Nhe*I and the insertion of a *Sal*I linker. The PCR template pGKVEA has a pGEM5Zf+ (Promega) backbone and between *Eag*I and *Apa*I restriction sites contains 2.5 kb of flanking sheep DNA, part of the tetracycline and neomycin resistance genes, and the first 575 nucleotides of the MVV_{KV1772} genomic RNA sequence. Mutagenesis (Quikchange, Stratagene) was carried out on plasmid pGVVS-N+S to generate plasmids pGTVVS and pGSVVS (see Figure 7). Primers for the former were as follows: forward, GAAAGAAAGCTTCGCGGACGCGTGAAGTAAGG, and reverse, CCTTACTTCACGCGTCCGCGAAGCTTTCTTTC. Primers for the latter were as follows: forward, GTGAAAGAAAGCTTCGGGTTCGCCTGAAGTAAGGTAAG, and reverse, CTTACCTTACTTCAGGCGAACCCGAAGCTTTCTTTCAC. Mutations were checked by sequence analysis.

RNA Production and Preparation. SP6 polymerase mediated transcription of *Sal*I linearized pGVVS-N+S generated sense orientated RNA from position 1 to position 480. Transcription templates, for use with the T7 polymerase Ribomax (Promega) system, were produced by Taq polymerase (Promega) mediated PCR using either pGKVEA or pBabepuro for the MoMLV (*17*) as a PCR template. Forward and reverse primers are described in Table 1. Freshly transcribed RNA was purified by ethanol precipitation and size exclusion chromatography through G-50 Sephadex columns (Roche) and then resuspended in diethyl pyrocarbonate (DEPC) treated water. Synthetic RNA oligonucleotides were purchased deprotected and desalted from Dharmacon RNA Technologies.

RNA Dimerization Assays. RNA was incubated for 1 h at 37 °C in dimerization buffer [2× buffer A = 100 mM cacodylic acid (pH 7.0), 80 mM KCl, and 10 mM $MgCl_2$; 2× buffer B = 100 mM cacodylic acid (pH 7.0), 1 M KCl, and 10 mM $MgCl_2$]. Where appropriate, the thermal treatment of the RNA before incubation is detailed in the figure legend. Following electrophoresis in 1× TBE (89 mM Tris−acetate, 89 mM boric acid, 2 mM EDTA) or 1× TBM (89 mM Tris−acetate, 89 mM boric acid, 0.7 mM magnesium chloride) samples were visualized by ethidium bromide staining.

For RNA heterodimerization experiments, RNA transcripts were denatured for 3 min at 80 °C and snap-cooled on ice. Equimolar RNAs were incubated for 1 h at 37 °C in dimerization buffer B (see above) before electrophoresis in 1× TBE and visualization by ethidium bromide staining. Quantification of relative levels of dimeric and monomeric RNA was performed using ImageJ v1.30 (Research Service Branch, NIH).

Antisense DNA Oligonucleotide Inhibition Assays. MVV1−485 (3 pmol) and a 50-fold molar excess of the appropriate DNA oligonucleotide were mixed in dimerization buffer B and denatured for 3 min at 80 °C. Samples were immediately placed at 37 °C and incubated for 1 h before electrophoresis in 1× TBM and ethidium bromide staining.

Computer Modeling and Phylogenetic Analysis. RNA secondary structures were mathematically modeled using Mfold and Pfold under the default conditions of 1 M NaCl and 37 °C (*18*−*20*). Nucleotide sequences were retrieved from the EMBL Nucleotide Sequence Database (Release 71) using their Sequence Retrieval System (SRS) and aligned using the online NClustalW function.

Transmission Electron Microscopy. Negative stain transmission electron microscopy was performed as described previously (*21*). In brief, RNA was resuspended in 5 mM $MgCl_2$, heated to 65 °C for 1 min, and cooled to room temperature over 30 min. RNA was diluted to 2.5 μg/mL, adsorbed onto a carbon film, and negatively stained with an aqueous uranyl acetate solution [4% (w/v), pH 4.25]. The adsorption time was 30 s followed by washing with doubly distilled water for 10 s and incubation on 4% uranyl acetate for 1 min. Specimens were mounted onto copper grids (400 mesh) and blotted dry. Micrographs were recorded on Agfa Scientia 23D56 electron image sheet films at calibrated magnifications in a Phillips CM10 transmission electron microscope operated at 100 kV.

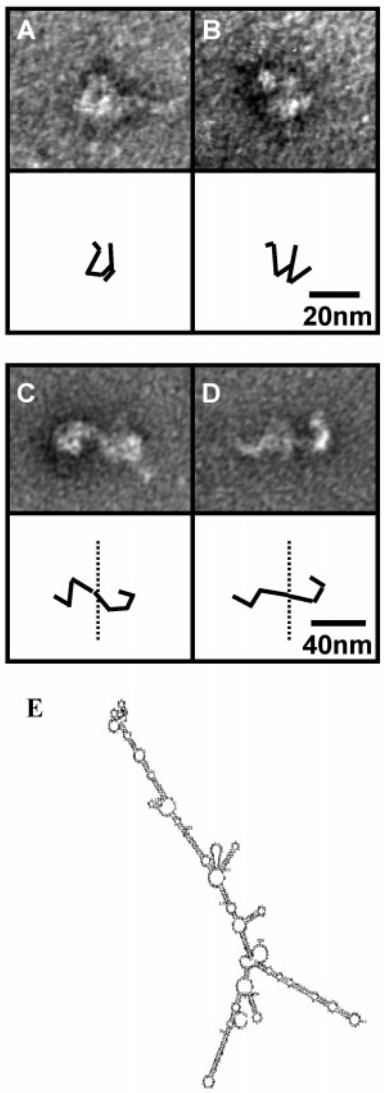

FIGURE 1: Global structure of the MVV leader region. (A−D) Negative stain transmission electron microscopy was performed on MVV1−485 as detailed in Experimental Procedures. Representative structures of monomeric (A and B) and putative dimeric (C and D) RNA are shown. The scale is as indicated on the individual figures. (E) Mfold of the MVV1−485 leader.

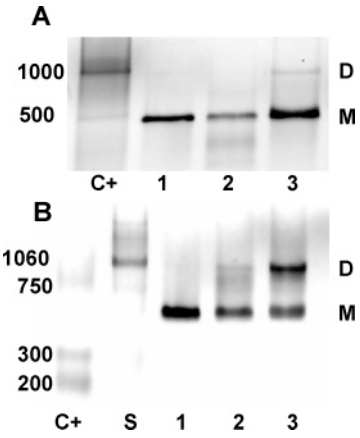

FIGURE 2: MVV RNA dimers require magnesium during electrophoresis. Standard dimerization assays were performed with 1.0 μg of sense MVV1−480 RNA. RNA was incubated for 1 h at 37 °C in either H_2O (lane 1), dimerization buffer A (lane 2), or dimerization buffer B (lane 3). Electrophoresis was performed using 1.8% Metaphor (Cambrex Bio Science) gels prepared with and run in either (A) TBE (89 mM Tris−borate, 2 mM EDTA) or (B) TBM (89 mM Tris−borate, 0.7 mM $MgCl_2$). C+ = Century Plus RNA markers (Ambion), S = dimeric HTLV RNA531 (a 531/1062 bp size marker), D = MVV dimer, and M = MVV monomer.

RESULTS

Investigating the Global RNA Conformation. Visualization of RNA dimers in vitro has been previously achieved by transmission electron microscopy (*22*). To investigate the global RNA conformation of the MVV leader RNA, and any alterations following dimerization, MVV1−485 was visualized by negative stain transmission electron microscopy. Figure 1A−D contains representative images of RNA structures corresponding to both monomer and putative dimer. Of 80 such molecules visualized, 73% appear to have formed dimers. The images suggest that monomeric MVV leader RNA does indeed form a compactly folded structure of 20−25 nm diameter formed from three stem−loops but with additional hinged folds in two of these (Figure 1A,B). The size of the monomeric RNA is consistent with that of previously visualized RNA structures of similar sequence length and proposed structural complexity (*23*). Importantly, these structures resemble closely the free energy model generated by Mfold (Figure 1E), in which three stem−loops can be seen, two of which contain additional smaller stem−loops. In contrast, the RNA dimer forms an extended linear structure of 60−70 nm with what is probably a single site of contact between the two RNA strands (Figure 1C,D). Formation of the dimer disrupts the compact folding seen for the monomer.

Isolating the MVV Dimerization Initiation Site. The in vitro assay established to visualize MVV RNA dimers revealed a requirement for magnesium ions as previously observed for HIV-1 (*24*). Incubation of MVV leader RNA, derived from the plasmid pGVVS-N+S, did not induce significant levels of dimer formation under conditions known to promote RNA dimerization of human T-cell lymphotropic virus type I (Figure 2A). However, replacement of the EDTA present during electrophoresis with physiological levels of $MgCl_2$ (0.7 mM) generated conditions conducive to MVV RNA dimerization, in which dimer formation was most pronounced at high cationic concentrations (Figure 2B). The formation of RNA dimers was maintained at higher magnesium concentrations of 1.4 and 5.0 mM but was not supported with only 0.1 mM magnesium (data not shown), suggesting a minimum requirement for divalent cations to stabilize structural interactions during dimer formation. KCl (1 M) was found to enhance dimerization as has recently been observed in HTLV-1 (*25*).

To isolate the core sequences involved in the dimerization interaction, the ability of truncated MVV RNA transcripts to maintain dimer formation was determined using standard dimerization assays. Representative results are provided and summarized in Figure 3.

Transcript MVV1−485, containing almost the entire leader sequence, demonstrated highly efficient dimer formation (88%), levels that were similar following transcript extension into the *gag* gene (MVV1−571, 71% dimer; Figure 3A). Truncation of the leader RNA to position 220 (MVV1−220) significantly reduced the level of the RNA dimer (12%). This was further diminished by removal of the PBS (MVV1−159, 8%), while the first 97 nucleotides of the leader region

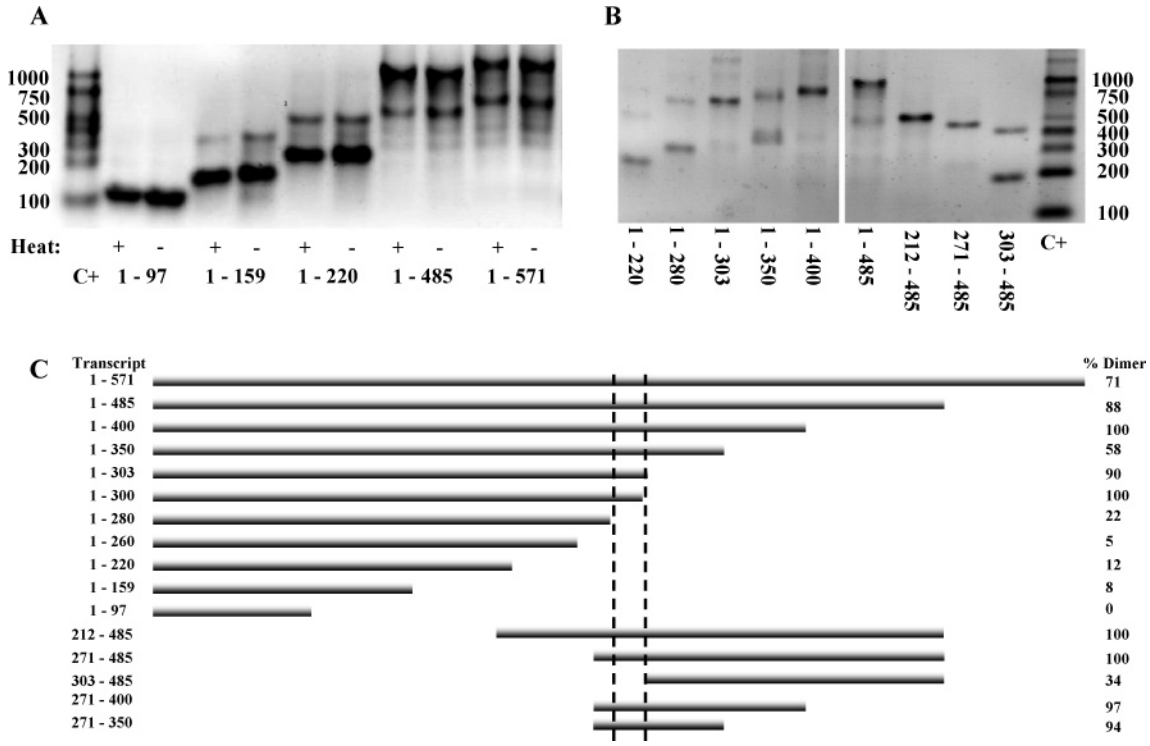

FIGURE 3: Dimerization characteristics of leader RNA truncations. (A) 1.0 μg of the RNA species described was treated in standard dimerization assays and electrophoresis performed in the presence of 0.7 mM magnesium. Samples were either thermally denatured for 3 min at 80 °C before incubation in dimerization buffer B (+) or incubated directly (−). (B) 1.0 μg of the RNA detailed was treated in standard dimerization assays without any prior thermal denaturation. C+ = Century Plus RNA markers. (C) Summary of the dimerization characteristics of the transcripts used in this study. Transcripts 1−260, 1−300, 271−400, and 271−350 are not included in the representative gels in (A) and (B).

(MVV1−97) were completely incapable of forming dimers. The dimeric character of these RNA transcripts was similar whether or not they were thermally denatured before incubation in dimerization buffer (compare paired samples in Figure 3A).

The dramatic change in levels of dimer formation between transcripts MVV1−485 and MVV1−220 suggested that the MVV DIS is located between positions 220 and 485. The dimerization characteristics of further truncations across this region confirmed this observation. The 3′ truncated RNA terminating at, or downstream of, position 280 remained predominantly monomeric (<22% dimer), while those terminating at, or upstream of, position 303 were generally over 80% dimeric (Figure 3B; compare MVV1−280 and MVV1−303). The sole exception to this was MVV1−350, which formed only 58% dimer. Similarly, truncations from the 5′ end lacking up to the first 271 nucleotides were more than 90% dimeric, while removal of the first 303 nucleotides reduced dimer formation to 34% (Figure 3B). While this suggests that a core DIS may reside between residues 281 and 303, both the upstream and downstream regions could potentially contribute to the stability of resultant dimers or enhance their formation in a multipartite manner.

Specific Inhibition of RNA Dimerization by Antisense DNA Oligonucleotides. To confirm the importance of the putative DIS for MVV RNA dimerization, the inhibitory effect on MVV1−485 RNA dimerization of a 50-fold molar excess of short antisense DNA oligonucleotides spanning nucleotides 203−400 was determined. Relative to the oligonucleotide free control sample every DNA oligonucleotide, except 303−281, promoted RNA MVV1−485 dimer formation,

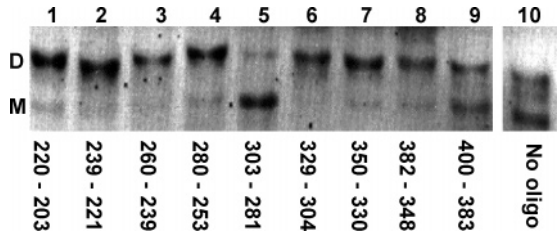

FIGURE 4: Antisense DNA oligonucleotides inhibit MVV RNA dimer formation. Dimerization inhibition assays were performed as described in Experimental Procedures. All lanes contain 3 pmol of MVV1−485 and 150 pmol of the antisense DNA oligonucleotide specified, except for lane 10 which only contains 3 pmol MVV1−485. D = MVV1−485 dimer and M = MVV1−485 monomer.

most likely as a result of increased nucleic acid concentration as is commonly observed with RNA (Figure 4). In contrast, dimerization was almost ablated by incubation with the DNA oligonucleotide 303−281 spanning the proposed core DIS domain (Figure 4, lane 5). This supports a critical role for this region in MVV RNA dimerization as suggested by the previous truncation studies.

A Stem−Loop Structure Is the Core Dimer Linkage. The leader regions of a total of eight MVV isolate sequences were aligned and compared. Computational analysis was performed using the MFold and Pfold algorithms in order to identify conserved RNA secondary structures. Conserved regions with identifiable structures were modeled (Figure 5); less conserved intervening sequences have been omitted. Where possible, putative roles were assigned to the motifs.

Within the domain 281−303 the RNA is predicted to fold into a stable GACG tetraloop closing an 8 base pair stem (Figure 6A). A synthetic RNA from position 280−301 can

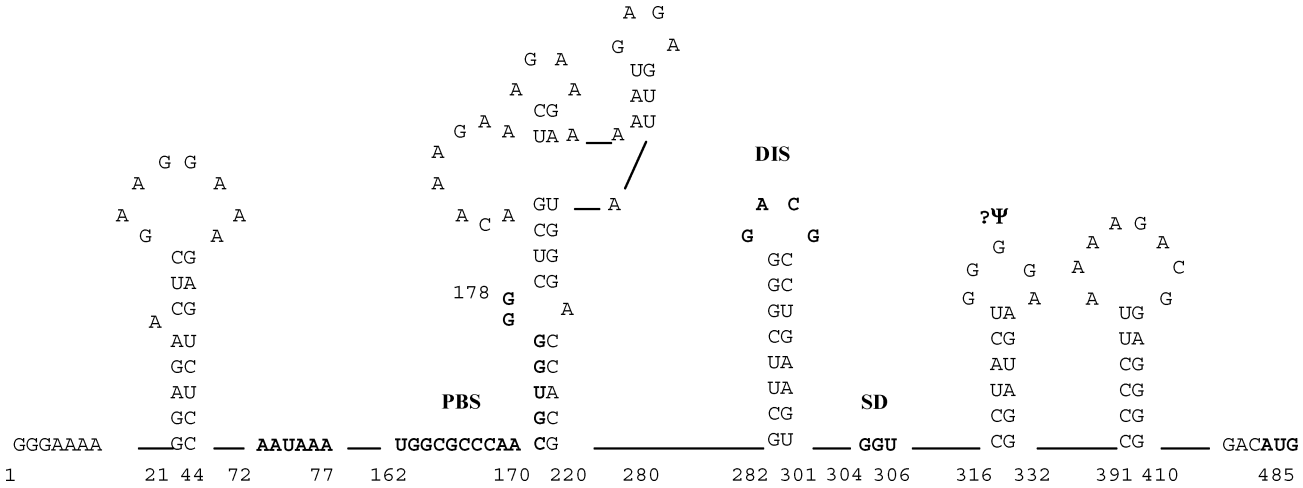

FIGURE 5: Proposed secondary structure of the maedi visna virus leader region. Conserved structures were determined by repeated MFold and Pfold analysis of eight (see Figure 8) MVV isolate sequences. Delineated in bold are putative regulatory elements. Numbering is according to ref *16*.

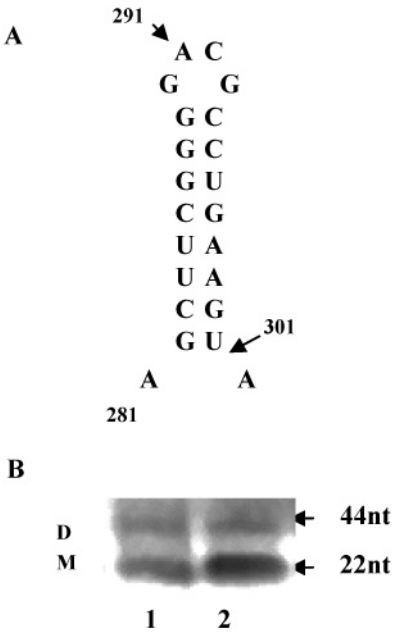

FIGURE 6: The MVV DIS is predicted to form a GACG tetraloop. (A) Mfold predicted structure of the putative DIS between positions 281 and 303 as identified by truncation and inhibition studies. The predicted free energy is −11 kcal/mol. (B) Synthetic RNAs of the MVV DIS and M-MLV H3 loop form dimers independently of accessory sequences. RNA is at 20 pmol/μL, diluted in 50 mM cacodylic acid, 1 M KCl, and 5 mM $MgCl_2$. Lanes: 1, MVV DIS, nucleotides 280−301, heated at 80 °C for 3 min, incubated on ice for 3 min, and then heated at 37 °C for 30 min; 2, M-MLV GACG H3 tetraloop. D = dimer and M = monomer.

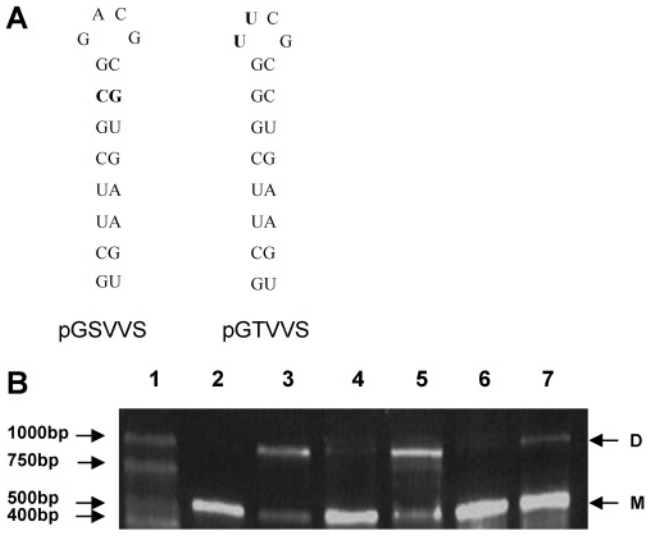

FIGURE 7: Mutations of the proposed DIS support the MFold predictions. (A) Mutations introduced into the MVV putative DIS sequence. Non-native bases are indicated in bold. (B) Lanes 2, 4, and 6: RNA (at 200 ng/μL) heated at 80 °C for 3 min and then incubated on ice. Lanes 3, 5, and 7: RNA heated and then incubated for 60 min at 37 °C. Lanes 2 and 3: wild-type RNA. Lanes 4 and 5: pGSVVS RNA. Lanes 6 and 7: pGTVVS RNA. Lane 1 = Century Plus RNA markers.

form dimers independently of surrounding RNA sequences (Figure 6B, lane 1), as does an identically sized RNA spanning the H3 loop from MLV (Figure 6B, lane 2). The lower levels of the RNA dimer observed suggest that while this motif may be capable of functioning as a DIS, this behavior is enhanced by the presence of additional leader sequences.

Mutations in the Presumed MVV DIS Support the Proposed Structure and Sequences as Important for Dimerization. The presence of four sequential guanines in the stem−loop raises the possibility that dimerization involves guanine quartets. To address this, pGSVVS containing a point mutation G287C was generated (Figure 7A). pGSVVS maintains the predicted helical region.

A second mutation, pGTVVS, was a two-base substitution in the terminal loop, changing the sequence from GACG to the well-studied and highly stable UUCG structure. The latter is known to have weak dimerization ability. Both of these mutations were assessed for dimerization in the context of the 1−485 leader sequence. pGSVVS did not alter the ability of the RNA to dimerize (Figure 7B, lane 5 versus lane 3). In both instances the RNA was dimeric. In contrast, RNA containing the tetraloop mutation, derived from plasmid pGTVVS, showed severely impaired dimerization (Figure 7B, lane 7). These data support the hypothesis that the loop sequences are those important for dimer formation, whereas some of the nucleotides predicted to form the helical stem are interchangeable. Data obtained from 1D and 2D NMR (carried out using the constructs shown in Figure 6B) are

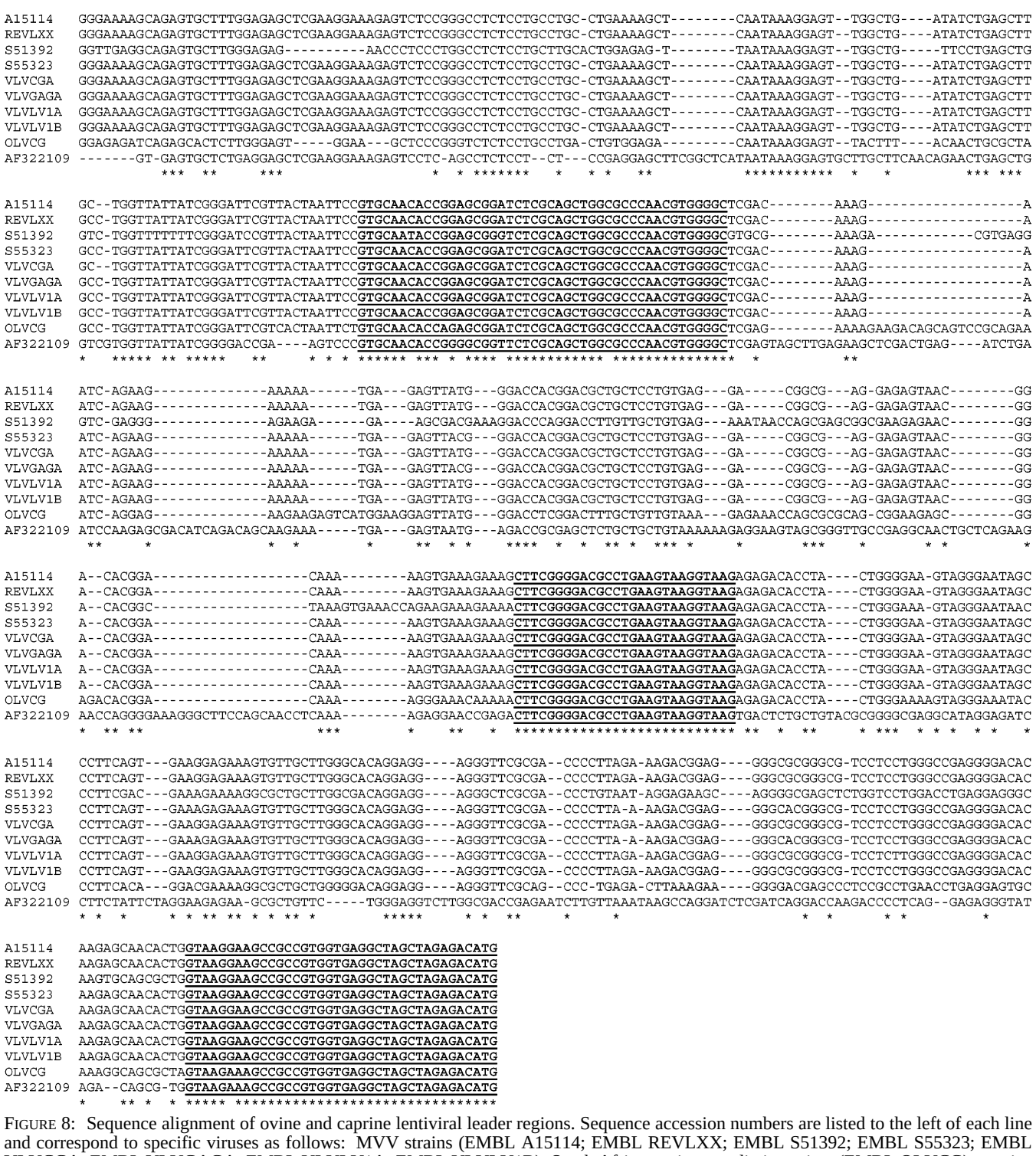

FIGURE 8: Sequence alignment of ovine and caprine lentiviral leader regions. Sequence accession numbers are listed to the left of each line and correspond to specific viruses as follows: MVV strains (EMBL A15114; EMBL REVLXX; EMBL S51392; EMBL S55323; EMBL VLVCGA; EMBL VLVGAGA; EMBL VLVLV1A; EMBL VLVLV1B); South African ovine maedi visna virus (EMBL OLVCG); caprine arthritis encephalitis virus (EMBL AF322109). Homology is identified by an asterisk. Regions of high identity are highlighted in bold and underlined. The positions of these regions (numbering from EMBL A15114) correspond to the viral primer binding site and immediate upstream region (134−178), the putative core DIS (280−300), and the region immediately before the *gag* coding sequence (454−491).

consistent with the loop−loop interaction previously found in the MoMLV H3 loop (Peter Lukavsky, personal communication).

The Putative MVV DIS Is Highly Conserved between Viral Isolates. Phylogenetic conservation of RNA sequences between viral isolates may be indicative of the formation of secondary structures with conserved functional roles (*26*). A comparative sequence alignment of the leader regions from the genomic sequences of eight MVV viral isolates, the divergent South African ovine maedi visna virus, and the related caprine arthritis encephalitis virus deposited in the EMBL Nucleotide Sequence Database (Release 71) identified three regions of significant sequence homology (Figure 8). These corresponded to the viral primer binding site and its immediate upstream sequence, the proposed DIS motif, and the 3′ termini of the leader RNA. The observed conservation of homology between the caprine and ovine lentiviruses supports an important functional role for the DIS stem loop.

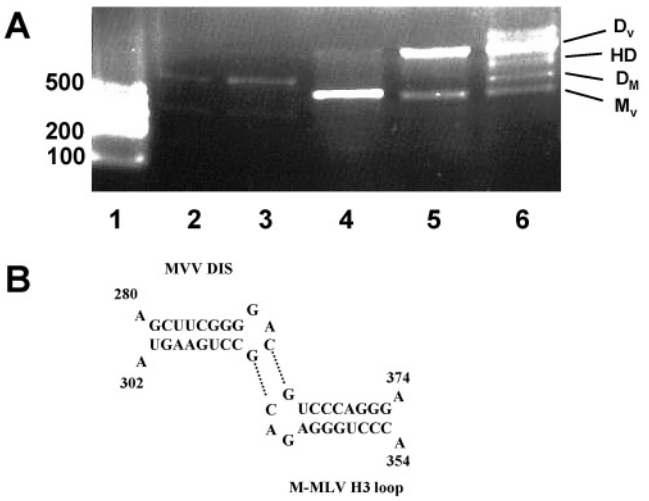

FIGURE 9: Formation of heterodimers between the MVV leader RNA and the M-MLV Ψ domain. (A) MVV leader RNA (485 bp) and M-MLV (331 bp) RNA were incubated either alone (lanes 4 and 5 and lanes 2 and 3, respectively) or at equimolar concentrations (lane 6) in dimerization buffer B. Transcripts were heated at 80 °C for 3 min to dissociate the dimers and then either incubated on ice (lanes 2 and 4) or incubated at 37 °C for 60 min (lanes, 3, 5, and 6). Lane 1 = Century Plus RNA markers. Ethidium bromide staining. D_M = M-MLV dimer, D_V = MVV dimer, HD = MVV:M-MLV heterodimer, M_M = M-MLV monomer, and M_V = MVV monomer. (B) Potential site of nucleic acid interactions between the MVV DIS and M-MLV H3 loop.

The MVV DIS Displays Structural Homology and Heterodimerizes with Motifs in the γ-Retrovirus Moloney Murine Leukemia Virus. The predicted GACG tetraloop forming the MVV DIS differs extensively from DIS motifs associated with primate lentiviral RNA dimerization. However, comparison with structural motifs in the leader regions of other retroviruses revealed surprising homology with motifs from the α- and γ-retroviruses. These viruses possess highly conserved GACG tetraloop motifs (*27*), believed to contribute to viral RNA packaging (*28*, *29*) and which can assist in the formation of RNA dimers (*30*). NMR and solvent molecular dynamic studies of the H3 GACG loop from MoMuLV have shown that these structures can dimerize via kissing interactions between the two 3′ terminal nucleotides of the tetraloop (*31*, *32*; Figure 6B, lane 2). Further, GACG tetraloops have also been shown to mediate heterodimer formation between Harvey sarcoma virus and murine leukemia viruses (*33*). To see if the MVV DIS motif could heterodimerize with the MoMuLV Ψ domain (a 331 bp transcript), the RNAs were incubated together at equimolar concentrations. A species consistent with a heterodimer was observed (Figure 9A, lane 6) and confirmed by ^{32}P end labeling of one of the RNA species (data not shown). The possible mechanism by which this interaction might occur is shown in Figure 9B. As shown in Figure 6B, the stem nucleotides differ between the MVV and MLV structures, indicating that the interaction is between the loop nucleotides.

DISCUSSION

The results presented in this paper detail the first in vitro study of the structure of the MVV leader and MVV RNA dimerization. The imaging and predictive modeling of the viral leader provide evidence for a major structural change in retroviral RNA between the monomeric and dimeric states. The MVV DIS is centered on a helix terminating in a GACG tetraloop between positions 281 and 300, a region which is highly conserved between the ovine and caprine lentiviruses. Intriguingly, this structure shows homology with structural motifs in the α- and γ-retroviruses but not with DIS regions identified in the primate lentiviruses. Given the structural and sequence homologies in the *Retroviridae*, this is most likely an example of convergent evolution, rather than reflecting a common origin, although a recombination event in the distant past is also a possibility.

The RNA truncation and DNA antisense inhibition studies both support the position of the MVV DIS as residing between nucleotides 281 and 300. While the sequences both upstream and downstream of this region are dispensable for dimer formation (Figures 3 and 4), it is conceivable that they may aid the formation of genomic dimers and increase dimer stability. Certainly various transcripts retain the ability to form low levels of dimer in the absence of the core DIS (e.g., MVV1−220 and MVV303−485; Figure 3). Some show reduced dimer formation even when retaining the DIS (MVV1−350; Figure 3) and may be required for efficient dimer formation. This behavior may relate to the presence of particular structural motifs or be an effect on global leader RNA folding and DIS presentation. It is also possible that the sequence from 303 to 350 has an inhibitory effect as has been noted in other viruses; however, this is overcome in the full-length transcript. The impact of such conformational variation would be comparable to recently modeled systems in HIV-1 and HIV-2 (*34*−*36*).

Further persuasive support for the presence of a critical structural, and probably also functional, motif between positions 281 and 300 is provided by the pattern of phylogenetic conservation across the leader regions of the ovine and caprine lentiviruses (Figure 8). The primer binding site, the putative DIS domain, and the 3′ terminus of the leader RNA are all highly conserved, suggesting critical functional roles for all of these regions.

The similarities between the proposed MVV DIS motif and RNA motifs in the leader regions of the α- and γ-retroviruses have already been noted. Within these viruses GACG tetraloops are involved in the packaging of viral RNA (*28*, *29*), and while not a component of the core M-MLV DIS motif (*37*), they may contribute to the process of dimerization and the stability of the resultant dimer (*30*). Of particular relevance to the dimerization of MVV RNA is the fact that the MoMuLV H3 GACG tetraloop has been shown by NMR and solvent molecular dynamic studies to be capable of forming a kissing interaction between the terminal CG residues of the tetraloop in each RNA strand (*31*, *32*). Equivalent interactions may occur between MVV DIS motifs leading to kissing hairpin formation and the possibility of subsequent maturation to an extended duplex dimer. Indeed, as mentioned above, preliminary NMR data confirm this. The initial interaction might be relatively unstable as it is likely to involve only two nucleotides in each loop and could thus account for the lower levels of dimer formed by the DIS motif in the absence of additional RNA sequences (Figure 6B). Further, this would be reconcilable with the observations of Brahic and Vigne (*15*) and with the requirement for magnesium to stabilize dimer formation (see Figure 2). Removal of stabilizing magnesium ions might facilitate rapid denaturation of the RNA dimer, and certainly the

susceptibility of immature, or weakly associating, retroviral RNA dimers to dissociate in the presence of EDTA as a result of magnesium chelation has been documented (*38*).

The specific formation of heterodimers between MVV and MoMuLV (Figure 9) raises parallels with recent studies of the dimerization of murine leukemia viruses and Harvey sarcoma virus in which GACG tetraloops were found to mediate interspecies RNA heterodimerization (*33*). However, the formation of stable MVV:M-MLV heterodimers was in many ways surprising, as the ability to form extended duplexes between the different RNAs would be significantly impaired due to the different nucleotide sequences of the RNA stem. Further, it is likely that the terminal CG nucleotides of the tetraloop, which are presumably involved in the dimerization interaction, have different stacking patterns for the two viruses as a result of the different terminal base pairs on the RNA stems, G_{288}•C_{293} for MVV and U_{328}•A_{333} and A_{362}•U_{367} for the M-MLV GACG tetraloops. This would imply a conformational rearrangement in at least one of the tetraloops to facilitate dimerization. If such rearrangement does occur, then it reinforces the strength and specificity of the kissing interactions utilized by retroviruses to initiate RNA dimer formation. This question is being addressed by our tertiary structural analysis of these structures.

ACKNOWLEDGMENT

The authors acknowledge the scientific input of the Sykes Trust. We also thank Teresa Wallman for secretarial assistance and Peter Lukavsky for useful discussion.

REFERENCES

1. Berkhout, B., and van Wamel, J. L. (1996) Role of the DIS hairpin in replication of human immunodeficiency virus type 1, *J. Virol. 70*, 6723−6732.
2. Ly, H., Nierlich, D. P., Olsen, J. C., and Kaplan, A. H. (2000) Functional characterization of the dimer linkage structure RNA of Moloney murine sarcoma virus, *J. Virol. 74*, 9937−9945.
3. McBride, M. S., and Panganiban, A. T. (1996) The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures, *J. Virol. 70*, 2963−2973.
4. Paillart, J. C., Marquet, R., Skripkin, E., Ehresmann, C., and Ehresmann, B. (1996) Dimerization of retroviral genomic RNAs: structural and functional implications, *Biochimie 78*, 639−653.
5. Paillart, J. C., Berthoux, L., Ottmann, M., Darlix, J. L., Marquet, R., Ehresmann, B., and Ehresmann, C. (1996) A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis, *J. Virol. 70*, 8348−8354.
6. Mikkelsen, J. G., Lund, A. H., Duch, M., and Pedersen, F. S. (2000) Mutations of the kissing-loop dimerization sequence influence the site specificity of murine leukemia virus recombination in vivo, *J. Virol. 74*, 600−610.
7. Clever, J. L., and Parslow, T. G. (1997) Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation, *J. Virol. 71*, 3407−3414.
8. Greatorex, J., and Lever, A. (1998) Retroviral RNA dimer linkage, *J. Gen. Virol. 79* (Part 12), 2877−2882.
9. Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994) Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro, *Proc. Natl. Acad. Sci. U.S.A. 91*, 4945−4949.
10. Girard, P. M., Bonnet-Mathoniere, B., Muriaux, D., and Paoletti, J. (1995) A short autocomplementary sequence in the 5′ leader region is responsible for dimerization of MoMuLV genomic RNA, *Biochemistry 34*, 9785−9794.
11. Laughrea, M., and Jette, L. (1996) Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248−271 are dispensable for dimer formation, *Biochemistry 35*, 1589−1598.
12. Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (1996) A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA, *Proc. Natl. Acad. Sci. U.S.A. 93*, 5572−5577.
13. Ly, H., Nierlich, D. P., Olsen, J. C., and Kaplan, A. H. (1999) Moloney murine sarcoma virus genomic RNAs dimerize via a two-step process: a concentration-dependent kissing-loop interaction is driven by initial contact between consecutive guanines, *J. Virol. 73*, 7255−7261.
14. Fu, W., and Rein, A. (1993) Maturation of dimeric viral RNA of Moloney murine leukemia virus, *J. Virol. 67*, 5443−5449.
15. Brahic, M., and Vigne, R. (1975) Properties of visna virus particles harvested at short time intervals: RNA content, infectivity and ultrastructure, *J. Virol. 15*, 1222−1230.
16. Andresson, O. S., Elser, J. E., Tobin, G. J., Greenwood, J. D., Gonda, M. A., Georgsson, G., Andresdottir, V., Benediktsdottir, E., Carlsdottir, H. M., and Mantyla, E. O. (1993) Nucleotide sequence and biological properties of a pathogenic proviral molecular clone of neurovirulent visna virus, *Virology 193*, 89−105.
17. Morgenstern, J. P., and Land, H. (1990) Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line, *Nucleic Acids Res. 18*, 3587−3596.
18. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol. 288*, 911−940.
19. Zuker, M., Mathews, D. H., and Turner, D. H. (1999) in *RNA Biochemistry and Biotechnology* (Clark, B. F. C., Ed.) pp 11−43, NATO ASI Series, Kluwer Academic Publishers, Dordrecht, The Netherlands.
20. Knudsen, B., and Hein, J. (2003) Pfold: RNA secondary structure prediction using stochastic context-free grammars, *Nucleic Acids Res. 31*, 3423−3428.
21. Beales, L. P., Rowlands, D. J., and Holzenburg, A. (2001) The internal ribosome entry site (IRES) of hepatitis C virus visualized by electron microscopy, *RNA 7*, 661−670.
22. Bender, W., Chien, Y. H., Chattopadhyay, S., Vogt, P. K., Gardner, M. B., and Davidson, N. (1978) High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures, *J. Virol. 25*, 888−896.
23. Beales, L. P., Holzenburg, A., and Rowlands, D. J. (2003) Viral internal ribosome entry site structures segregate into two distinct morphologies, *J. Virol. 77*, 6574−6579.
24. Marquet, R., Paillart, J. C., Skripkin, E., Ehresmann, C., and Ehresmann, B. (1994) Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site, *Nucleic Acids Res. 22*, 145−151.
25. Monie, T. P., Greatorex, J. S., Zacharias, M., and Lever, A. M. L. (2004) The human T-cell lymphotropic virus type-1 dimerization initiation site forms a hairpin loop unlike previously characterised retroviral dimerization motifs, *Biochemistry 43*, 6085−6090.
26. Mathews, D. H., and Turner, D. H. (2002) Dynalign: an algorithm for finding the secondary structure common to two RNA sequences, *J. Mol. Biol. 317*, 191−203.
27. Konings, D. A., Nash, M. A., Maizel, J. V., and Arlinghaus, R. B. (1992) Novel GACG-hairpin pair motif in the 5′ untranslated region of type C retroviruses related to murine leukemia virus, *J. Virol. 66*, 632−640.
28. Mougel, M., and Barklis, E. (1997) A role for two hairpin structures as a core RNA encapsidation signal in murine leukemia virus virions, *J. Virol. 71*, 8061−8065.
29. Fisher, J., and Goff, S. P. (1998) Mutational analysis of stem-loops in the RNA packaging signal of the Moloney murine leukemia virus, *Virology 244*, 133−145.
30. De Tapia, M., Metzler, V., Mougel, M., Ehresmann, B., and Ehresmann, C. (1998) Dimerization of MoMuLV genomic RNA: redefinition of the role of the palindromic stem-loop H1 (278−303) and new roles for stem-loops H2 (310−352) and H3 (355−374), *Biochemistry 37*, 6077−6085.

31. Kim, C. H., and Tinoco, I., Jr. (2000) A retroviral RNA kissing complex containing only two G·C base pairs, *Proc. Natl. Acad. Sci. U.S.A. 97*, 9396−9401.
32. Reblova, K., Spackova, N., Sponer, J. E., Koca, J., and Sponer, J. (2003) Molecular dynamics simulations of RNA kissing-loop motifs reveal structural dynamics and formation of cation-binding pockets, *Nucleic Acids Res. 31*, 6942−6952.
33. Rasmussen, S. V., Mikkelsen, J. G., and Pedersen, F. S. (2002) Modulation of homo- and heterodimerization of Harvey sarcoma virus RNA by GACG tetraloops and point mutations in palindromic sequences, *J. Mol. Biol. 323*, 613−628.
34. Huthoff, H., and Berkhout, B. (2002) Multiple secondary structure rearrangements during HIV-1 RNA dimerization, *Biochemistry 41*, 10439−10445.
35. Huthoff, H., and Berkhout, B. (2001) Two alternating structures of the HIV-1 leader RNA, *RNA 7*, 143−157.
36. Dirac, A. M., Huthoff, H., Kjems, J., and Berkhout, B. (2002) Regulated HIV-2 RNA dimerization by means of alternative RNA conformations, *Nucleic Acids Res. 30*, 2647−2655.
37. Ly, H., and Parslow, T. G. (2002) Bipartite signal for genomic RNA dimerization in Moloney murine leukemia virus, *J. Virol. 76*, 3135−3144.
38. Lanchy, J. M., and Lodmell, J. S. (2002) Alternate usage of two dimerization initiation sites in HIV-2 viral RNA in vitro, *J. Mol. Biol. 319*, 637−648.

BI048529M