JOURNAL OF VIROLOGY, May 1996, p. 2963–2973 Vol. 70, No. 5
0022-538X/96/$04.00+0

The Human Immunodeficiency Virus Type 1 Encapsidation Site Is a Multipartite RNA Element Composed of Functional Hairpin Structures

M. SCOTT MCBRIDE AND ANTONITO T. PANGANIBAN*

McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, Madison, Wisconsin 53706

Received 21 September 1995/Accepted 1 February 1996

We analyzed the leader region of human immunodeficiency virus type 1 (HIV-1) RNA to decipher the nature of the *cis*-acting E/Ψ element required for encapsidation of viral RNA into virus particles. Our data indicate that, for RNA encapsidation, there are at least two functional subregions in the leader region. One subregion is located at a position immediately proximal to the major splice donor, and the second is located between the splice donor and the beginning of the *gag* gene. This suggests that at least two discrete *cis*-acting elements are recognition signals for encapsidation. To determine whether specific putative RNA secondary structures serve as the signal(s) for encapsidation, we constructed primary base substitution mutations that would be expected to destabilize these potential structures and second-site compensatory mutations that would restore secondary structure. Analysis of these mutants allowed the identification of two discrete hairpins that facilitate RNA encapsidation in vivo. Thus, the HIV-1 E/Ψ region is a multipartite element composed of specific and functional RNA secondary structures. Compensation of the primary mutations by the second-site mutations could not be attained in *trans*. This indicates that interstrand base pairing between these two stem regions within the hairpins does not appear to be the basis for HIV-1 RNA dimer formation. Comparison of the hypothetical RNA secondary structures from 10 replication-competent HIV-1 strains suggests that a subset of the hydrogen-bonded base pairs within the stems of the hairpins is likely to be required for function in *cis*.

Encapsidation of genomic viral RNA during virion assembly is an essential step in the retrovirus replication cycle. During this process, full-length genomic viral RNA is incorporated into the nascent virus particle and subgenomic viral RNA is largely excluded (38). On the basis of work with a variety of retrovirus species, the process of RNA encapsidation can be considered as the recognition of *cis*-acting elements resident on the viral RNA by one or more virus-encoded proteins. The nucleocapsid protein (NC) is derived from proteolytic processing of the Gag precursor protein, and NC, or the NC domain of the Gag precursor protein, is directly or indirectly required in *trans* for RNA encapsidation (4, 6, 11, 12, 16, 20, 43, 44, 53). However, it is not clear whether NC is sufficient for specific recognition of full-length retroviral RNA. Current in vitro data are consistent with a role for additional Gag-derived peptides in the encapsidation process in the case of bovine leukosis virus (29).

For some retroviruses, the *cis*-acting element necessary for encapsidation, interchangeably referred to as either E or Ψ (E/Ψ), is located in the untranslated leader region between the major subgenomic splice donor and the *gag* start codon (37, 41, 58). Such placement of the encapsidation signal at a site distal to the major subgenomic splice donor provides the virus with the means to specifically recognize and encapsidate full-length genomic viral RNA and to exclude spliced RNA species; spliced RNAs would not be encapsidated efficiently simply because they lack the E/Ψ site (40). The location of the avian leukosis virus (ALV) E/Ψ site is a notable exception to this spatial mechanism by which full-length and spliced RNAs may be differentiated. In ALV, the encapsidation site lies upstream of the splice donor, so the E/Ψ site would be expected to be found on both full-length and spliced RNAs (6, 7, 30, 34, 50, 54). The exact mechanism by which ALV discriminates between full-length and spliced RNAs has not yet been elucidated. In addition to the primary encapsidation site, additional regions play ancillary roles in the encapsidation process. These include the 5′ end of the *gag* gene and a region near the 3′ end of the viral RNA (2, 5, 10, 56). These regions of the RNA facilitate encapsidation measurably only in conjunction with E/Ψ. Since they do not exhibit function in the absence of E/Ψ, these regions are considered to be of secondary importance in encapsidation. In contrast, E/Ψ can be sufficient to direct significant encapsidation.

Based on the precedent with simple retroviruses, mutational analysis of the region between the major splice donor and the *gag* gene has been carried out previously to determine whether that region of human immunodeficiency virus type 1 (HIV-1) RNA is required for encapsidation (4, 15, 25, 36). Although initial data indicated that the region between the splice donor and *gag* is profoundly important in encapsidation, subsequent analyses have suggested that alternative regions are important for encapsidation (11–13, 16, 31, 38, 48, 51). One of these regions encompasses the 5′ end of *gag*, and the second is located proximal to the major subgenomic splice donor.

Conserved secondary structures within the 5′ untranslated region have been suggested for a variety of retroviruses (6, 23, 32). Specific RNA secondary structures necessary for encapsidation of spleen necrosis virus (SNV) have been previously identified by genetic analysis in vivo (59). In addition, the presence of specific secondary structures within the presumptive encapsidation region of HIV-1 RNA has been hypothesized on the basis of biochemical-enzymatic analysis, phylogenetic comparisons, and computer modeling (9, 16, 24, 52). Moreover, some of these potential secondary structures appear to facilitate binding of synthetic HIV-1 RNAs to Gag protein

* Corresponding author. Mailing address: McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, 1400 University Ave., Madison, WI 53706. Phone: (608) 263-7820. Fax: (608) 262-2824. Electronic mail address: Panganiban@oncology.wisc.edu.

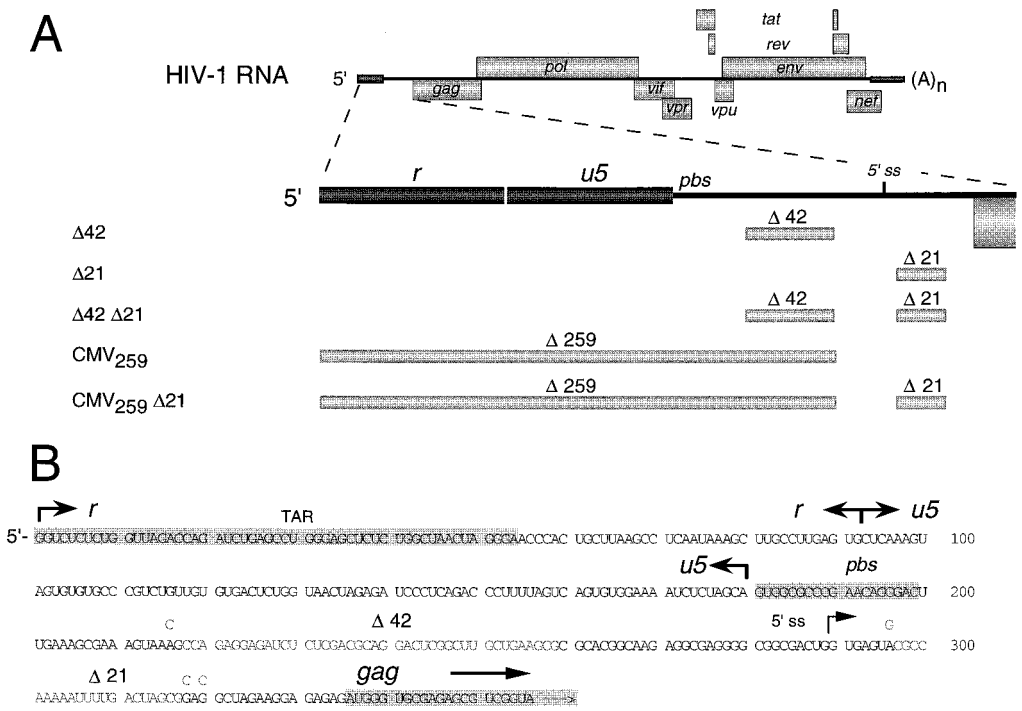

FIG. 1. Nature of deletion mutations in the leader region of the HIV-1 genome. (A) The spatial organization of the 5′ end of the viral genome, shown in expanded view. Significant landmarks in this region include the 5′ copy of the terminal repeat region (r), the unique 5′ sequence (u5), the primer binding site for minus-strand DNA synthesis (pbs), the principal splice donor for the formation of the singly and multiply spliced mRNA species (5′ ss), and the beginning of the *gag* gene. The locations of the 42-, 21-, and 259-nt deletions (Δ42, Δ21, and Δ259) are indicated. In $pCMV_{259}$ and $pCMV_{259}$Δ21, a human CMV promoter-enhancer is used for expression. (B) Nucleotide sequence of the untranslated leader from HIV_{NL4-3} and location of some associated *cis*-acting elements. The Tat-responsive element (TAR), pbs, and the beginning of *gag* (shaded boxes) and nucleotides that have been removed by deletion in the Δ42 and Δ21 mutations (pale letters) are indicated. The Δ42 mutation contains two base substitution mutations immediately 5′ to the 42-nt deletion. Similarly, the Δ21 mutation contains three substitution mutations adjacent to the 21-nt deletion. These base substitution mutations are indicated above the original sequence.

in vitro (11, 12, 16, 53). Further, RNA pseudotyping experiments, in which RNA from simian immunodeficiency virus is efficiently incorporated into HIV-1 particles, indicate that higher-order RNA structure is likely to be required in *cis* for encapsidation (52). We report here that a specific hairpin structure proximal to the major subgenomic splice donor functions in encapsidation in vivo. However, a second hairpin distal to the splice donor is also an encapsidation signal. Interestingly, disruption of either of these hairpins alone by deletion or base substitution has only a partial effect on encapsidation. However, the simultaneous ablation of both has a profound effect on encapsidation efficiency. Thus, the HIV-1 encapsidation region is a multipartite element and contains functional and distinct encapsidation elements.

MATERIALS AND METHODS

Construction of plasmid DNAs. Standard techniques were used for molecular cloning (39). pMSMBA is a derivative of the infectious proviral clone pNL4-3 (1). For the purposes of describing pMSMBA and related mutant plasmids, nucleotide designations refer to the DNA sequence of pNL4-3. To facilitate the generation of mutants, we first constructed a plasmid, pMSMBA, which contains the entire viral component of pNL4-3 but is devoid of extraneous cellular DNA in pNL4-3. The long terminal repeat (LTR) of pNL4-3 was amplified by PCR with a sense primer complementary to the sequence of nucleotides (nt) 1 to 21 (5′-TGGAAGGGCTAATTTGGTCCC) (primer 1-21) and an antisense primer complementary to the sequence of nt 634 to 612 (5′-TGCTAGAGATTTTCCA CACTGAC) (primer 634-612). The amplified LTR was cloned into plasmid pT7 (Novagen) to create pT7LTR. pT7LTR was digested with *Spe*I and *Eco*RI, and the LTR-containing fragment was subcloned into the *Xba*I and *Eco*RI sites in pGEM11zf(−) (Promega) to create pGEMLTR. pGEMLTR was digested with *Bsp*EI, and the coding sequence of pNL4-3 from the *Bsp*EI site at nt 308 to the *Bsp*EI site at nt 9384 was inserted to create $pGEM_{NL4-3}$. An in-frame stop codon at nt 6359 and a deletion from nt 6365 to 7252 were created within *env* by PCR amplification with sense primer 5762-5783 (5′-GTTTATCCATTTCAGAATT GGG) and mismatch antisense primer 6374-6352 *Nhe*I (5′-GAGTGGTG CTAGCTTACTTCCAC). The latter primer contains a G→T transversion at nt 6359, resulting in the introduction of an in-frame stop codon within *env*, as well as an *Nhe*I site at nt 6362. The amplified 612-bp fragment was digested with *Sal*I and *Nhe*I and cloned into the corresponding *Sal*I and *Nhe*I sites in $pGEM_{NL4-3}$ to create pMSMΔEnv2. To facilitate the introduction of mutations into the leader region, the *Bsp*EI site at nt 9383 was converted to an *Age*I site by the ligation of two fragments obtained by PCR amplification of pMSMΔEnv2 using two sets of primers, sense primer 8856-8879 (5′-GAGCCAGCAGCAGAT GGGGTGGGA) with antisense primer 9397-9374 *Age*I (5′-TTGAAGTACAC CGGTTGCAGCTCT) and sense primer 9374-9397 *Age*I (5′-AGAGCTGC AACCGGTGTACTTCAA) with antisense primer T7 (5′-TGTAATACGACT CACTATAGGGCG), complementary to the T7 promoter in pGEM11zf(−). After amplification, the fragments obtained from the 8856-8879 and 9397-9374 *Age*I primer pair and the 9374-9397 *Age*I–T7 primer pair were digested with *Age*I and ligated. The ligation mixture was amplified with sense primer 8856-8879 and antisense primer T7 and digested with *Xho*I and *Sfi*I to generate a 900-bp fragment containing the substitution mutations. This fragment was ligated to pMSMΔEnv2 digested with *Xho*I and *Sfi*I to create pMSMBA.

pΔ42 was created by PCR amplification of a fragment of pMSMBA using sense primer 275-298 (5′-CCTAGCATTTCGTCACATGGCCCG) and mismatch antisense primer 688-660 *Mlu*I (5′-CGAGAGATCTCCTCACGCGTT ACTTTCGC) to create an *Mlu*I site at nt 671. The resulting 413-bp fragment was digested with *Bsp*EI and *Mlu*I and ligated to pMSMBA digested with *Bsp*EI and *Bss*HII to create a deletion from nt 672 to 713 and a base substitution at nt 670 (Fig. 1). pΔ21 was created by the ligation of two fragments obtained by PCR amplification of pMSMBA with two sets of primers, the sense primer 275-298 with the antisense mismatch primer 763-739 *Apa*LI (5′-AAAATTTTTGGT GCACTCACCAGTC) and the sense mismatch primer 760-784 *Apa*LI (5′-TTTT GACTAGTGCACGCTAGAAGGA) with the antisense primer 1535-1511 (5′-CATCCTATTTGTTCCTGAAGGGTAC). The two fragments were digested with *Apa*LI and ligated. The ligation mixture was then amplified with the sense primer 275-298 and the antisense primer 1535-1511 to generate a 1,239-bp fragment containing a 21-bp deletion from nt 751 to 772 and base substitutions at nt 740, 772, and 774 (Fig. 1). This fragment was digested with *Bsp*EI and *Spe*I and ligated to pMSMBA digested with *Bsp*EI and *Spe*I. pΔ42Δ21 was created by

the same strategy used to create pΔ21 but with pΔ42 as the template for the PCR amplification.

Cloning strategies similar to that used for generation of pΔ21 were used to create pS1, pS2, pS3, pS3′, and pS4 using the mismatch primer pairs 687-715 *Eco*RI (5′-CGACGCAGGACTCGGAATTCTGAAGCGCG) with 716-692 *Eco*RI (5′-GCGCGCTTCAGAATTCCGAGTCCTG), 743-774 *Age*I (5′-GGT GAGTAACCGGTAAATTTTG) with 774-743 *Age*I (5′-CAAAATTTACCGGT TACTCACC), 764-789 *Mlu*I (5′-GACTAGCGGAGACGCGTAGGAGAGA G) with 789-766 *Mlu*I (5′-CTCTCTCCTACGCGTCTCCGCTAG), 756-779 *Mlu*I (5′-AAAATTTTGACGCGTGGAGGCTAG) with 779-756 *Mlu*I (5′-CTAGCCTCCACGCGTCAAAATTTT), and 785-809 *Apa*I (5′-GAGAGATG GGGGCCCGAGCGTCGGT) with 809-785 *Apa*I (5′-ACCGACGCTCGGGC CCCCATCTCTC), respectively, in the PCR amplification. pS1S2, pS1S3, pS2S3, pS1S2S3, pS1S3S4, pS1Δ21, and pΔ42S3 were created by using the same cloning strategy with the appropriate primers and templates for PCR amplification.

pS1/S1′S3 and pS1/S1′S3′ were created by the same strategy with pS3 or pS3′ used as a template for the PCR amplification, respectively, by the mismatch primer pair 687-734 *Eco*RI (5′-CGACGCAGGACTCGGAATTCTGAAGCGC GCACGGAATTAGGCGAG) and 729-687 *Eco*RI (5′-CGCCTAATTCCGTG CGCGCTTCAGAATTCCGAGTCC). pS1′S3′ was created by substituting the *Bsp*EI-*Bss*HII fragment of pMSMBA into the corresponding *Bsp*EI and *Bss*HII sites in pS1/S1′S3′.

pS1S3′/S3 was created by cloning of the *Not*I-*Spe*I fragment (nt −14 to 1506) of pS1 into pBluescript SK+ (Stratagene) followed by oligonucleotide-mediated site-directed mutagenesis of the single-stranded DNA by a standard protocol (39). In that mutagenesis, we used the antisense oligonucleotide 789-761 *Mlu*I (5′-CTCTCTCCTACGCGTCTCCACGCGTCAAA). The resulting mutation was transferred back to pMSMBA by substitution with a *Bsp*EI-*Spe*I fragment (nt 308 to 1506). pS1/S1′S3′/S3 was created by the strategy described above using pS1S3′/S3 as a template for the PCR amplification with the mismatch primer pair 687-734 *Eco*RI and 729-687 *Eco*RI.

pMSM50 was created by cloning the *Eco*RI-*Sph*I fragment of p763 (provided by Bill Sugden, University of Wisconsin—Madison) into pBS$^-$ (Stratagene). The *Xba*I-*Xho*I fragment of KT81 (52) was cloned into pMSM50 to create $pMSM50_{NL4-3}$. The *Bss*HII-*Sma*I fragment of pMSMBA or pΔ21 was cloned in $pMSM50_{NL4-3}$ digested with *Bss*HII and *Nae*I to create $pCMV_{259}$ or $pCMV_{259}\Delta21$, respectively. $pCMV_{-30}$ was created by PCR amplification of $pMSM50_{NL4-3}$ using the primer pair cytomegalovirus (CMV) 1-12 *Not*I (5′-CTATAGGGCCGCTCAT GTTTG) and HIV432/CMV789 (5′-CTTATATGTAGACCTCCCAC) and PCR amplification of pMSMBA using the primer pair CMV789/HIV432 (5′-GTGGG AGGTCTACATATAAG) and 789-761 *Mlu*I. The two fragments were mixed and amplified with CMV 1-12 *Not*I and 789-761 *Mlu*I. The resulting fragment was digested with *Not*I and *Bss*HII and ligated into pMSMBA.

pGEM(600-900) was created by PCR amplification of pMSMBA using the primer pair 587-610 *Bam*HI (5′-ACTAGAGATGGATCCGACCCTTTT) and 911-888 *Xho*I (5′-AGCTCCCTGCTCGAGCATACTATA). The 324-nt fragment was digested with *Bam*HI and *Xho*I and ligated into pGEM11zf(−).

Mutations were verified by sequence analysis.

Transfections. Twenty-four hours before transfection, HeLa cells were seeded at a density of 1.5×10^6/100-mm-diameter plate in Dulbecco modified Eagle medium supplemented with 7% calf serum and incubated at 37°C in 5% CO_2. Transfection was carried out with 3 to 10 μg of plasmid DNA per 100-mm-diameter plate by the calcium phosphate precipitation method (3) in the presence of chloroquine (Sigma) (60 μg/ml) for 6 to 8 h followed by a 15% glycerol shock for 4 min. The cells were then washed with phosphate-buffered saline and fed with fresh medium. 293 cells were transfected in a similar manner except that they were seeded at a density of 3.5×10^6/100-mm-diameter plate in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum, 2× BES (*N*,*N*-bis[2-hydroxyethyl]-2-aminoethane sulfonic acid) was substituted for 2× HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid), and chloroquine was omitted.

Isolation of RNA from cytoplasm and virions. After 48 to 72 h, the media and cytoplasmic RNA were concurrently collected. The cytoplasmic RNA was harvested as previously described (39), and the concentration was determined by spectrophotometric absorption at a λ of 260 nm. Virus was harvested from the media by first removing cellular debris by centrifugation at 4,000 rpm in a Beckman AccuspinFR table-top centrifuge. The virus was then pelleted from the clarified media by centrifugation through a 20% sucrose cushion for 2.5 h at 25,000 rpm in an SW28 rotor. The viral pellet was resuspended in TNE (10 mM Tris-HCl [pH 7.4], 100 mM NaCl, 1 mM EDTA). The physical virus titer was determined by using an antigen-capture assay to quantitate p24 (Coulter Cytometry), and RNA was isolated from equivalent p24 units for each experiment. To isolate RNA, virions were disrupted by the addition of sodium dodecyl sulfate (SDS) to 1% and incubation in a boiling water bath for 5 min. The sample was then treated with proteinase K (500 μg/ml) at 37°C for 30 min and subjected to phenol (pH 4.5)-chloroform extraction, chloroform extraction, and ethanol precipitation in the presence of 20 μg of poly(C) RNA. The isolated nucleic acid was resuspended in 50 μl of a buffer containing 10 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, and 10 U of RNase-free DNase I (Boehringer Mannheim) and incubated at 37°C for 30 min.

Nucleic acid analysis. RNA slot blot hybridization analysis was performed as described previously (39). An *Apa*LI-*Msc*I fragment from $pCMV_{259}\Delta21$, spanning pNL4-3 nt 748 to 4551, was used to generate a probe by random priming in the presence of [α-^{32}P]dCTP. Hybridization was carried out in the presence of Rapid-Hyb buffer (Amersham). Filters were washed extensively with 0.2× SSC (1× SSC is 150 mM NaCl and 15 mM sodium citrate [pH 7.0])–0.1% SDS at 65°C. Signals were detected and quantitated by phosphorimage analysis (Molecular Dynamics).

RNase protection. The antisense probe (1.2×10^8 cpm/μg) was synthesized by transcription of pGEM(600-900) with T7 RNA polymerase (Promega, Madison, Wis.) according to a previously described protocol (39) following linearization with *Not*I. To serve as size markers for denaturing polyacrylamide gels, ^{32}P-end-labeled fragments of pGEM11zf(−) digested with *Hpa*II were synthesized. Either 25 μg of cytoplasmic RNA or the virion RNA preparation from 10 ml of supernatant supplemented with 25 μg of poly(C) RNA was mixed with 5×10^5 Cerenkov counts of ^{32}P-labeled antisense RNA (3 fmol) and precipitated with ethanol. Samples were resuspended in 20 μl of hybridization buffer (100 mM sodium citrate [pH 6.4], 300 mM sodium acetate [pH 6.4], 1 mM EDTA in 80% formamide), heated at 95°C for 3 min, and hybridized at 42°C for 16 h. A 200-μl volume of RNase digestion mixture (10 mM Tris-HCl [pH 7.4], 300 mM NaCl, 5 mM EDTA, 2 μg of RNase T_1 per ml, 5 μg of RNase A per ml) was added for a 30-min incubation at 37°C. SDS was added to a final concentration of 1%, and proteinase K was added to 0.5 mg/ml. Samples were incubated for 30 min at 37°C, phenol-chloroform extracted, chloroform extracted, and precipitated with ethanol in the presence of 25 μg of poly(C) RNA carrier. Pellets were dissolved in 10 μl of formamide loading buffer (2 mM EDTA [pH 8.0] in 80% formamide), heated at 95°C for 3 min, and subjected to polyacrylamide gel electrophoresis (6% acrylamide/bis 19:1; 8 M urea). The various protected RNA species were quantitated by phosphorimage analysis (Molecular Dynamics).

RESULTS

Multiple functional elements within E/Ψ. To locate the RNA sequences within the untranslated leader region that are involved in HIV-1 encapsidation, we introduced a variety of mutations into the appropriate segment of a plasmid encoding nearly full-length viral RNA. pMSMBA contains both a point mutation, which creates a premature stop codon, and a deletion within the *env* gene. The remainder of the virus genome remains intact. Since Env protein is not required for virus production or release of particles from cells, transfection of cells with pMSMBA results in the production of virion particles that can be used to measure the efficiency of RNA encapsidation. pMSMBA is competent for one cycle of replication when complemented with a functional *env* gene (data not shown), and the efficiency of nucleic acid replication is similar to that of other previously studied molecular clones that lack a functional *env* gene (19, 26, 46, 49). This indicates that pMSMBA contains the *cis* elements necessary for replication including encapsidation.

The major subgenomic splice donor, which is used for the genesis of both singly and multiply spliced RNAs, is located 289 nt from the 5′ end of the viral RNA. Initially, we constructed a 42-nt deletion mutation just upstream of the splice donor (corresponding to RNA nt 218 to 259; Fig. 1) and a separate 21-nt deletion that removes a segment between the splice donor and the *gag* initiation codon (nt 297 to 317 in the viral RNA; Fig. 1). The resulting deletions were designated pΔ42 and pΔ21, respectively. We then attempted to determine the effect of these mutations on viral RNA encapsidation. Thus, HeLa cells were transfected in parallel with pMSMBA, pΔ42, and pΔ21. Two days later, virus was harvested and purified from the media while cytoplasmic RNA was isolated from the transfected cells. Viral RNA was then extracted from equivalent amounts of virus, as determined by an antigen-capture assay to detect the capsid (CA) protein, p24 (see Materials and Methods). This virion RNA, along with the cytoplasmic viral RNA, was then detected by hybridization analysis using a *gag*-specific probe, and the RNA was quantified by phosphorimage analysis. This evaluation of pΔ42 and pΔ21 indicated that neither mutation significantly affected steady-state levels of cytoplasmic RNA (Fig. 2A) or particle production, as evidenced by measurement of p24 units in the virus

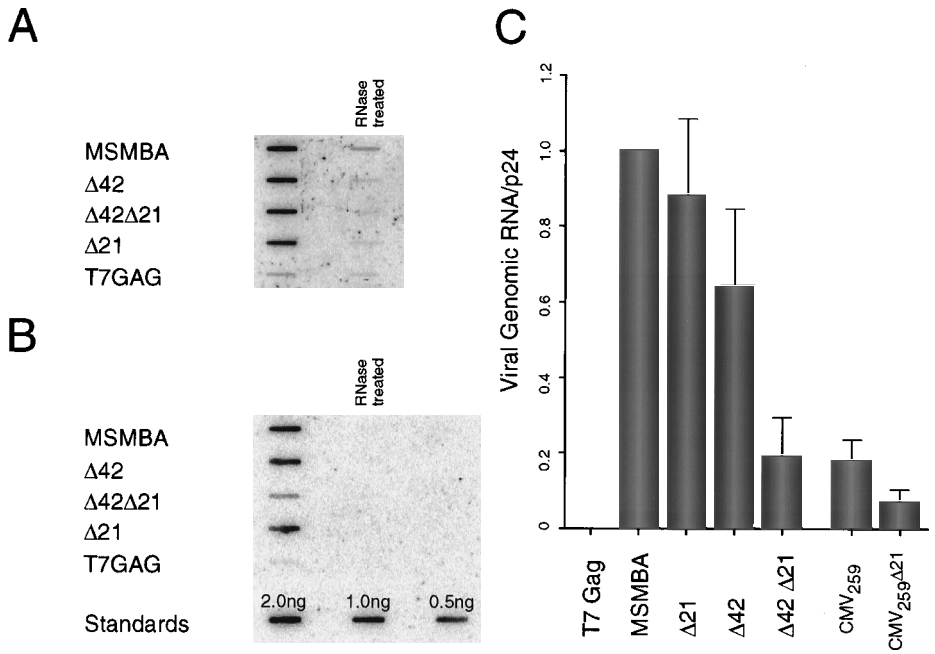

FIG. 2. Encapsidation of RNA containing deletion mutations in the untranslated leader region of HIV_{NL4-3}. Intracellular (cytoplasmic) and virion-associated HIV-1 RNA was detected and quantified as described in the text. pMSMBA contains a wild-type leader region. The nature of the Δ42 and Δ21 mutations is indicated in Fig. 1. pT7GAG (provided by M. Schwartz, University of Wisconsin) is a plasmid that contains the *gag* gene but lacks a eukaryotic promoter-enhancer. As such, the *gag* gene is not expressed following transfection, and pT7GAG serves as a control for plasmid DNA contamination. (A and B) Representative RNA slot blots used to detect viral RNA. (A) Equivalent amounts of cytoplasmic RNA from transfected cells were transferred to nitrocellulose, and genomic viral RNA was detected by using a *gag*-specific probe. Parallel samples were treated with RNase A prior to application to the filter. (B) RNA from equal amounts of virus (equivalent p24 units) was detected with a *gag*-specific probe. (C) Encapsidation efficiency. The relative amount of RNA from virus particles was quantified by phosphorimage analysis. The value of the wild-type control was normalized. The data are derived from five or more independent experiments.

preparation (data not shown). Moreover, neither mutation substantially reduced RNA encapsidation relative to that of the wild-type virus (Fig. 2B and C, pMSMBA). In the case of pΔ21, encapsidation was similar to that of the wild type, and for pΔ42 encapsidation was reduced marginally.

The observation that deletion of the region either immediately upstream or downstream of the splice donor resulted in only a modest reduction in encapsidation efficiency is consistent with two possibilities. Either the nucleotides removed by deletion might not be required for encapsidation, or there could be *cis* elements on either side of the splice donor that can independently promote encapsidation. To distinguish between these two alternatives, we introduced the two mutations together in *cis* into a single viral mutant designated pΔ42Δ21 (Fig. 1). In addition, to assess the general contribution that the entire region between the 5′ end of the RNA and the splice donor makes to encapsidation, we deleted all viral sequences upstream of nt 259 and introduced a human CMV promoter-enhancer to attain transcriptional expression of the remaining viral nucleic acid. This plasmid was designated $pCMV_{259}$ (Fig. 1). Finally, we constructed a related plasmid in which the 21-nt deletion between the splice donor and *gag* was introduced into $pCMV_{259}$ (designated $pCMV_{259}\Delta21$) (Fig. 1).

Quantitative hybridization and phosphorimage analysis of these mutants indicated that all mutants expressed similar levels of steady-state cytoplasmic RNA relative to the wild type (pMSMBA), and all produced similar amounts of virus following transfection, as evidenced by comparison of absolute p24 units (data not shown). However, in contrast to pΔ42 and pΔ21, the mutant bearing lesions both upstream and downstream of the splice donor, pΔ42Δ21, exhibited a significant reduction in encapsidation efficiency (Fig. 2B and C). These data are consistent with the possibility that the sequences immediately adjacent to the splice donor, including the primary RNA sequences and higher-order structures affected by the 42- and 21-nt deletions, both contribute significantly to encapsidation in *cis*.

Encapsidation of $pCMV_{259}$ was reduced in efficiency to a level similar to that of pΔ42Δ21 (Fig. 2). The region of the viral genome removed in $pCMV_{259}$ encompasses the nucleotides removed in the 42-bp deletion of pΔ42. Since the Δ42 mutation alone has little effect on encapsidation, it is likely that there is yet another element upstream of the 42-nt deletion that also contributes to encapsidation. Finally, $pCMV_{259}\Delta21$ exhibited an even more substantial reduction in encapsidation efficiency, suggesting that, in a general sense, the untranslated leader region is absolutely required for efficient encapsidation.

The fact that the 42- and 21-nt deletions did not affect intracellular genomic RNA levels or virus production indicated that neither mutation abrogated splicing from the adjacent splice donor; significant expression and cytoplasmic accumulation of full-length viral RNA would have required expression of Tat and Rev, which are translated from spliced mRNAs. However, as described below, we carried out experiments to determine whether mutations within the regions removed in the 42- and 21-nt deletions affected the efficiency of splicing. Moreover, we attempted to ascertain whether mutation of these regions affected the encapsidation of spliced RNA relative to full-length RNA.

We performed an ancillary control experiment to determine whether the substitution of the HIV LTR with the CMV promoter-enhancer might have had some unforeseen effect on splicing or encapsidation. We created an expression construct in which the entire HIV leader, beginning at the 5′ end of the

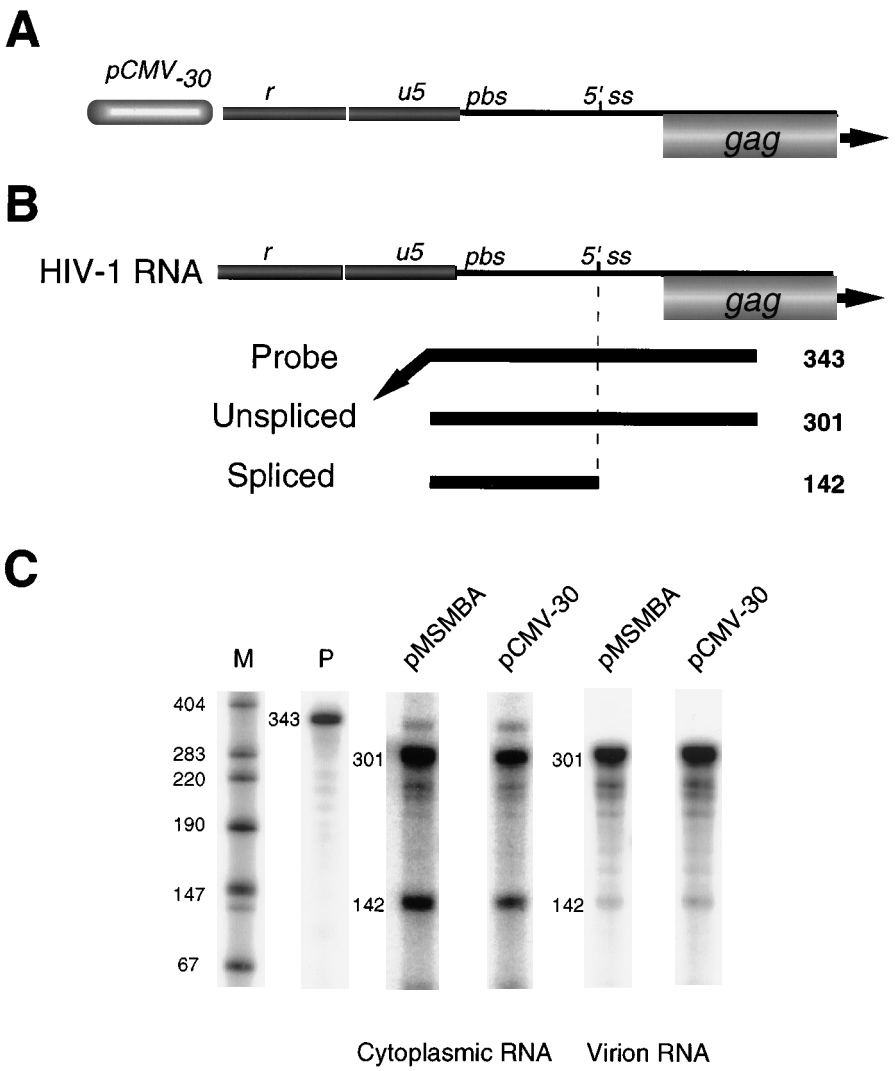

FIG. 3. RNase protection to detect encapsidation of HIV-1 RNA expressed from the CMV promoter-enhancer. (A) Regions of complementarity between the riboprobe generated from pGEM(600-900) and viral RNA. The probe is 343 nt long, and the diagnostic fragments for genomic and subgenomic RNAs are 301 and 142 nt, respectively. Abbreviations are as defined in the legend to Fig. 1. (B) Full-length HIV-1 RNA is expressed from the CMV promoter-enhancer in pCMV$_{-30}$. (C) Equivalent amounts of RNA were isolated from the cytoplasm of HeLa cells transiently transfected with pMSMBA or pCMV$_{-30}$, or RNA was isolated from equivalent physical titers of virus (equivalent p24 units) and subjected to RNase protection analysis using the pGEM(600-900) probe. The relative encapsidation efficiency of spliced RNA to unspliced RNA is approximately 10%. Fragment sizes (in nucleotides) are shown to the left of the marker and probe lanes M and P.

terminal repeat region, was expressed from the CMV promoter-enhancer. That plasmid was designated pCMV$_{-30}$ (Fig. 3A). To compare the expression, splicing, and encapsidation of RNAs derived from the CMV promoter-enhancer with those of RNAs derived from the HIV LTR, we transfected pCMV$_{-30}$ into cells in parallel with pMSMBA (Fig. 3B). Following the preparation of viral and cytoplasmic RNA, we analyzed the intracellular and viral RNAs, using an RNase protection assay. The probe in the RNase protection assay corresponded to the region encompassing the viral splice donor, so it was possible to identify full-length and spliced RNAs (Fig. 3C). The results of this experiment indicated that gene expression, splicing, and encapsidation were all equivalent for RNA expressed either from the LTR or from the CMV promoter-enhancer. The relative encapsidation efficiency of spliced RNA, the ratio of spliced versus unspliced RNAs in the virion relative to the ratio of these two RNAs in the cytoplasm, is approximately 10%.

In summary, the results from the encapsidation experiments with these various mutants were consistent with the presence of multiple discrete elements in the HIV-1 leader that contribute to RNA encapsidation. Since the Δ42 and Δ21 mutations suggested a small, defined portion of the leader region for further scrutiny, we chose to determine more directly whether there were *cis*-acting encapsidation elements lost as a result of the 42-nt and the 21-nt deletions and, if so, to elucidate the nature of those signals.

Specific secondary structures within HIV-1 RNA are encapsidation signals. Secondary structure has been proposed for the untranslated leader region of HIV-1 by biochemical-enzymatic analysis, phylogenetic comparisons, and computer modeling (9, 16, 24, 52), but little of this potential secondary structure has been tested genetically through base substitutions and second-site mutations designed to restore base-pairing within stems. The Δ42 and Δ21 mutations are expected to directly disrupt or remove hairpins 1, 2, and 3 (Fig. 4). To investigate

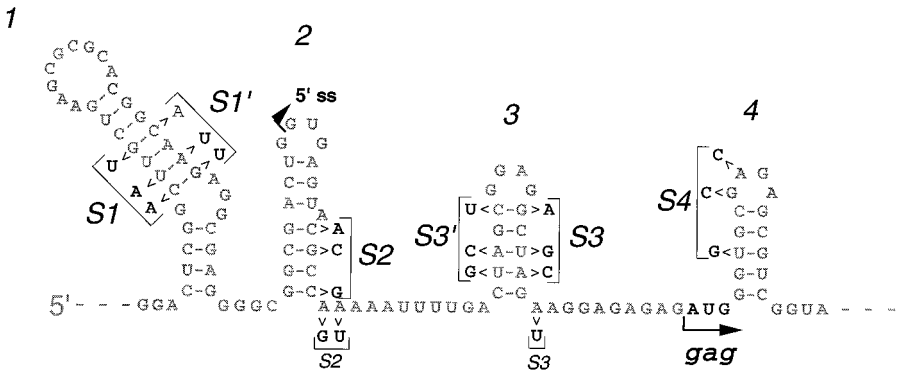

FIG. 4. Base substitution mutations in the leader region of HIV_{NL4-3}. The putative RNA secondary structure, based on previous analysis with other HIV strains closely related to HIV_{NL4-3} (1), and the theoretical structure of HIV_{NL4-3} RNA itself based on Mfold (University of Wisconsin Computer Genetics Group) (27, 28, 60) are shown. At least two alternative numerical designations for some of these hairpins have been used previously (16, 24). The designations presented in the figure are arbitrary and simply intended for clarity of presentation of the data. Substitution mutations that would destabilize the stems of these hairpins and second-site mutations that would be likely to restore the stems are shown adjacent to the wild-type sequence. Evaluation of likely secondary structures for each of the mutants by Mfold indicated that the substitutions designed to destabilize a particular stem of a hairpin would be unlikely to indirectly affect the formation of the other three hairpins. Mutation Δ42 (Fig. 1) contains a deletion that removes all the nucleotides that would constitute the 5′ side of the stem of hairpin 1. The deletion in mutation Δ21 would result in loss of the nucleotides that would constitute the 3′ side of the stem of hairpin 2 and the 5′ side of the stem of hairpin 3. The major splice donor (5′ ss) and the start codon of the *gag* gene are indicated.

the role of these structures in encapsidation, we created base substitution mutants designed to individually disrupt the intramolecular base pairing of each of the stems within the hairpins (pS1, pS2, and pS3) (Fig. 4). It has been previously reported that a region spanning the *gag* start codon is required for encapsidation (38). Moreover, there is a potentially stable hairpin structure located at this site, arbitrarily designated structure 4 here. To compare the role of this region with that of the region affected by the Δ42 and Δ21 mutations, we also constructed base substitution mutations within the stem of this hairpin (pS4) (Fig. 4). The mutations created in stem 4 are silent mutations with regard to amino acid sequence, but since there is a U-to-G transversion in the center of the 5-nt-long stem, formation of hairpin 4 would be highly unlikely.

In concordance with the results we obtained with the deletion mutants, which suggested that the regions immediately upstream and downstream of the splice donor both contribute to encapsidation in *cis*, singular mutations either upstream or downstream of the splice donor that separately affect stem 1, 2, or 3 had only a marginal effect on encapsidation (Fig. 5). In contrast, a double mutant, pS1S3, containing base substitutions both upstream and downstream of the splice donor, designed to simultaneously disrupt stems 1 and 3, was impaired in encapsidation to a level similar to that of pΔ42Δ21. In contrast, ablation of stem 2 did not reduce encapsidation efficiency. In fact, the base substitutions in mutant pS2 appeared to slightly increase encapsidation efficiency, as evidenced by comparison of the encapsidation efficiency of pS1S2 and pS1S2S3 with that of pS1 and pS1S3, respectively. Disruption of stem 4 (pS4) slightly reduced encapsidation efficiency in a manner similar to that of mutations that individually affect structure 1 or 3. However, a triple mutation that disrupts structures 1, 3, and 4 (pS1S3S4) did not reduce encapsidation efficiency below that of pS1S3.

To further verify that the primary mutations in stems 1 and 3 might account for the phenotype of the double mutant pΔ42Δ21, we also introduced the base substitution mutations S1 and S3 into constructs in combination with the deletion mutations Δ21 and Δ42, creating pS1Δ21 and pΔ42S3, respectively. Analysis of these mutants indicated that combination of Δ21 with S1 and Δ42 with S3 resulted in viruses that are reduced in encapsidation efficiency to a level similar to those of pΔ42Δ21 and pS1S3 (Fig. 5). Thus, the base substitution mutations and the deletion mutations produced similar phenotypes. Moreover, the results with all combinations of deletion and base substitution mutations were consistent with the presence of two discrete elements that both facilitate encapsidation.

To test the hypothesis that hairpins 1 and 3 are encapsidation elements, we created mutants containing the original stem mutations along with second-site mutations that potentially restore the secondary structure of stems 1 and 3 (Fig. 4 and 6). The mutations in pS1/S1′S3 should restore stem 1 (stem 1′) while stem 3 should remain disrupted. The mutations in pS1S3′/S3 should restore stem 3 (stem 3′) while stem 1 should remain disrupted. Finally, the mutations in pS1/S1′S3′/S3 should simultaneously restore the overall structure of both stem 1 and stem 3 (stem 1′ and stem 3′). The theoretical decrease in free energy that would result from formation of

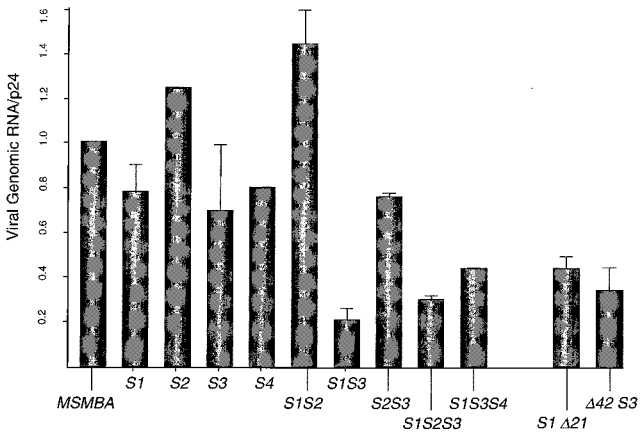

FIG. 5. Encapsidation efficiency of base substitution mutations. The efficiency of encapsidation was measured as for Fig. 2. The relative amounts of RNA from equivalent physical titers of virus (equivalent p24 units) were determined by using hybridization and phosphorimage analysis. pMSMBA is wild type, and the nature of each of the mutations analyzed is presented in Fig. 1 and 3. The data are derived from at least three independent experiments with the exceptions of data for pS2, pS4, and pS1S3S4.

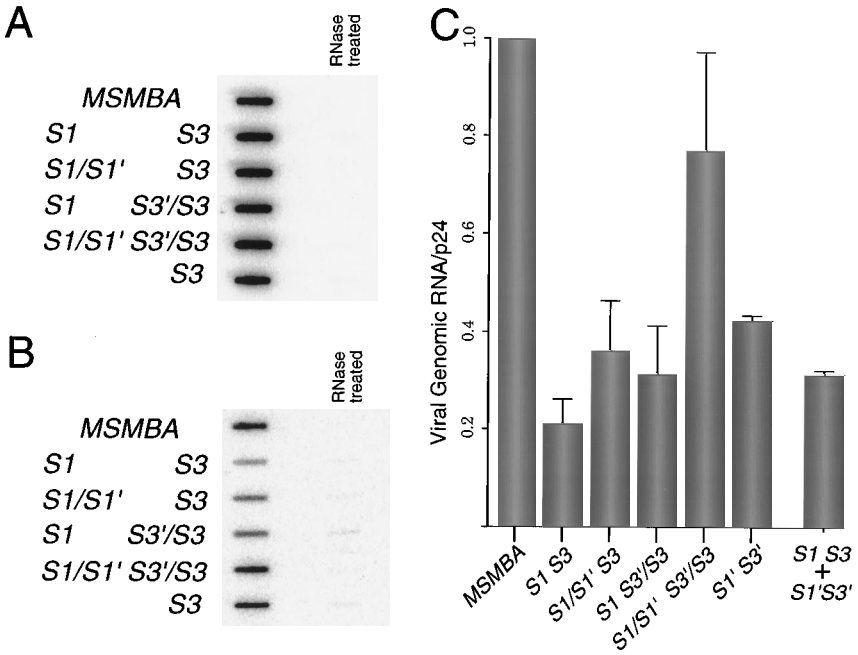

FIG. 6. Encapsidation efficiency of substitution mutations designed to restore secondary structure. (A) Cytoplasmic RNA; (B) virion RNA; (C) efficiency of encapsidation, measured as for Fig. 2. The indicated substitution mutations are described in the text and in the legend to Fig. 4.

both of these mutant stem-and-loop structures, stems 1′ and 3′, is similar to that expected for the formation of their parental wild-type counterparts, stems 1 and 3 (data not shown), on the basis of free energy calculations from Mfold (University of Wisconsin Genetics Computer Group) (27, 28, 60). We next examined the encapsidation efficiency of these second-site mutants directly. All of the mutants expressed similar levels of steady-state viral cytoplasmic RNA relative to the wild type (pMSMBA) (Fig. 6A). Significantly, the second-site mutations partially restored encapsidation alone (Fig. 6B and C, pS1/S1′S3 and pS1S3′/S3 and together (Fig. 6B and C, pS1/S1′S3′/S3) in comparison with the primary mutations expected to abrogate hydrogen bonding. These data indicate that hairpins 1 and 3 function as *cis*-acting encapsidation signals. However, it is important that restoration of encapsidation efficiency is incomplete.

Although the most straightforward interpretation of the data was that hydrogen bonding and stem formation are required for the formation and function of hairpins 1 and 3, a formal alternative was that the second-site compensatory mutations, S1′ and S3′, augment RNA encapsidation autonomously. To test this possibility directly, the effects of the S1′ and S3′ mutations were examined in the absence of the S1 and S3 mutations (pS1′S3′). As expected, pS1′S3′ exhibited diminished RNA encapsidation relative to that of pMSMBA or pS1/S1′S3′/S3 (Fig. 6).

Retroviral dimeric RNAs are noncovalently but stably associated with each other in virus particles. On the basis of thermodynamic considerations, the formation of intramolecular hairpins would be more likely than the alternative, intermolecular hairpins formed between the two RNAs within the dimer. However, it is possible that through a catalyzed or spontaneous interaction, the nucleotides that constitute the 5′ side of the stem within a hairpin might base pair with the 3′ side of the corresponding stem within a hairpin on the second RNA. In such a scenario, compensation of the S1 mutations by the S1′ mutations and compensation of the S3 mutations by the S3′ mutations could occur through intermolecular interaction and heterodimer formation. Consequently, we attempted to determine whether the observed phenotypic reversion could be manifested through such a *trans* interaction. pS1S3 and pS1′S3′ were coexpressed in cells, and we measured the encapsidation efficiency of virions produced by these cells. The results of this experiment indicated that the encapsidation efficiency was similar to that of either mutant alone (Fig. 6C). Thus, phenotypic reversion observed with the second-site mutations was likely due to restoration of secondary structure in stems 1 and 3 in *cis* rather than from interaction between two separate RNAs.

We next used an RNase protection assay to try to verify the data which indicated that hairpins 1 and 3 are both functional encapsidation signals. pS1, pS3, pS1S3, and pS1/S1′S3′/S3 were each separately cotransfected into cells along with a wild-type viral expression construct (pMSMBA). By using the same probe that was used in the RNase protection assay of Fig. 3, it was possible to measure the level of cytoplasmic viral RNA, evaluate the relative efficiency of splicing, and also determine the effect of each of the mutations on relative encapsidation efficiency. The expression and splicing in each of these mutants were equivalent (Fig. 7). Moreover, ablation of stems 1 and 3, in pS1, pS3, and pS1S3, resulted in a decrease in the encapsidation of full-length viral RNA (0.36, 0.25, and 0.20, respectively), and restoration of those two hairpin loops, in pS1/S1′S3′/S3, resulted in increased encapsidation efficiency (0.56). Removal of most of the 5′ end of the RNA in $pCMV_{259}$ and $pCMV_{259}\Delta 21$ resulted in the most severe decrease in encapsidation efficiency (0.04 and 0.02, respectively), underscoring the importance of this region of HIV-1 RNA in encapsidation.

In the cotransfections of pMSMBA with mutants that disrupt stem 3 (pS3 and pS1S3), a two- to threefold increase in the relative encapsidation of spliced RNA versus unspliced pMSMBA RNA is observed compared with the relative encapsidation efficiency of spliced RNA in cells transfected with pMSMBA alone. This increase is not observed upon cotrans-

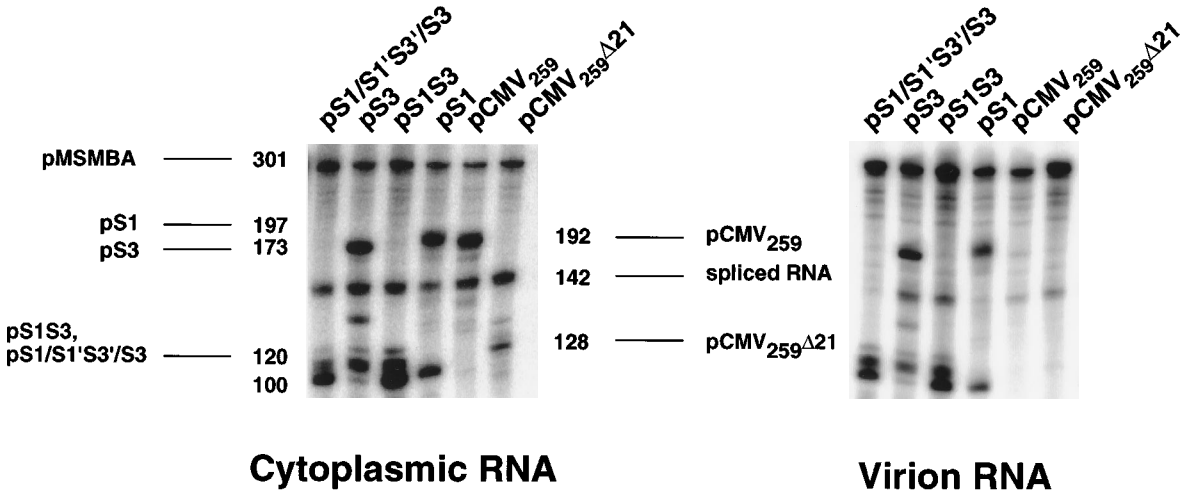

FIG. 7. Packaging of spliced and unspliced viral RNAs in HIV-1 E/Ψ mutants. pMSMBA was cotransfected with the indicated mutants. RNase protection analyses of the virion RNAs or the cytoplasmic RNAs were performed, using the pGEM(600-900) probe. The sizes (in nucleotides) and positions of diagnostic bands for genome-length RNAs from pMSMBA, pS1, pS3, pS1S3, pS1/S1′S3′/S3, $pCMV_{259}$, and $pCMV_{259}\Delta21$ are indicated. The diagnostic band for spliced RNA from pMSMBA and pS3 (142 nt) is also indicated. The 100-nt band is diagnostic for both unspliced and spliced RNAs from pS1, pS1S3, and pS1/S1′S3′/S3. Relative encapsidation efficiencies were calculated as the ratio of the amount of mutant to wild-type (pMSMBA) genomic RNAs in the virions, with normalization to the cytoplasmic levels of the two RNAs. Relative encapsidation efficiencies: pS1, 0.36; pS3, 0.25; pS1S3, 0.20; pS1/S1′S3′/S3, 0.56; $pCMV_{-30}$, 0.04; $pCMV_{-30}\Delta21$, 0.02.

fection of pMSMBA with pS1. These data are consistent with a specific role for stem 3 in packaging of full-length viral RNA and exclusion of spliced mRNA, since mutations within stem 3 decreased the efficiency of discrimination between full-length and spliced RNAs in the cotransfections. The second-site mutations within pS1/S1′S3′/S3 restore the ability to discriminate between full-length and spliced RNAs (Fig. 7).

If hairpins 1 and 3 are important for encapsidation, then it is likely that the same or similar structures would be maintained for efficient replication in diverse HIV-1 strains. Previous alignment of the leader region of available HIV sequences coupled with examination of putative secondary structure has proved to be a valuable approach to identifying possible conserved secondary structure in the E/Ψ region (16, 24). We tried to extend this theoretical analysis further by incorporating three additional criteria. First, previous comparisons reasonably relied on the battery of all available HIV-1 sequences. However, many of those sequences were from replication-defective virus strains, and we wanted to compare the sequences only of strains that are known to be capable of replication; we used nine other replication-competent HIV-1 strains for comparison with $HIV_{NL4-3/NY5}$. In addition to several members of the clade B group, to which HIV_{NL4-3} belongs, more-diverse sequences were also used in this comparison. The most disparate sequence was that of HIV_{cpz}, which shares only about 83% nucleotide identity with the sequence of HIV_{NL4-3}. Second, we used a recently devised program called Mfold-Phylo (University of Wisconsin Genetics Computer Group) to simultaneously compare the potential secondary structures of these multiple HIV-1 strains. Mfold-Phylo is a powerful way to specifically evaluate individual nucleotide pairs that may be hydrogen bonded with each other in related RNA molecules; regions involved in potential stem regions are evaluated more stringently than with Mfold. Third, we displayed the "p-num" analysis (27, 28, 60) from Mfold-Phylo to indicate the relative likelihood that particular base-paired and single-stranded regions form in each virus. The p-num value of a nucleotide reflects the number of potential pairing partners for that nucleotide; the predicted hydrogen bonding of nucleotides with lower p-num values is more likely to be significant.

The overall results of this RNA secondary structure comparison are presented in Fig. 8. The results of Mfold-Phylo indicate that the double-stranded regions of hairpins 1, 3, and 4 can potentially be found in all of the replication-proficient viruses examined. However, not all strains would be able to form the upper part of stem 2. p-num analysis of these virus strains also indicates a high level of confidence for the double- and single-stranded regions beginning just upstream of stem 3 and extending through stem 4. In contrast, the region comprising hairpins 1 and 2 has a slightly greater number of alternative potential pairing partners. However, compared with the nucleotides immediately upstream of stem 1 and distal to stem 4 (not shown), the nucleotides in this region have remarkably low p-num values. Taken together with the mutational analysis of the individual elements, it is likely that each of the replication-competent viruses can form hairpins 1, 3, and 4 and that these elements function in encapsidation in diverse HIV-1 strains.

DISCUSSION

On the basis of primary base substitution mutations and phenotypic reversion by second-site mutations, at least two hairpins (1 and 3) function in HIV-1 RNA encapsidation. Furthermore, although these hairpins show no apparent primary sequence similarity, they both appear to contribute to encapsidation in *cis*. Either hairpin 1 or hairpin 3 must be present for efficient encapsidation; simultaneous ablation of hairpins 1 and 3, via base substitution or deletion, substantially reduces encapsidation efficiency. It is of interest that the hairpins restored by second-site mutations do not function quite as well as their wild-type counterparts. This suggests that primary sequence could contribute to the role of hairpins 1 and 3 in E/Ψ function or that the mutations designed to restore secondary structure still affect a specific tertiary interaction required for maximal E/Ψ function.

Concurrent disruption of hairpins 1 and 3 reduces encapsidation efficiency to levels considered to represent a deficiency in RNA encapsidation in simple retroviruses. However, it is noteworthy that detectable encapsidation still occurs in the

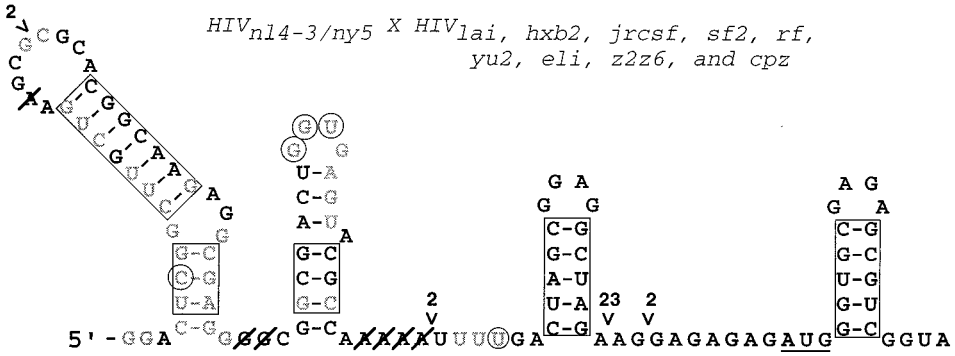

FIG. 8. Comparison of RNA secondary structure in E/Ψ region of 10 replication-competent HIV-1 strains. Mfold-Phylo (University of Wisconsin Genetics Computer Group) (27, 28, 60) was used to simultaneously examine the secondary structure of multiple aligned HIV-1 primary nucleic acid sequences. Although abundant sequence data for numerous HIV-1 isolates are available, only molecular clones that have been reported to be proficient for production of infectious virus following transfection were selected for this comparison. Primary sequences were aligned by using the Genetics Computer Group program Pileup. Gaps in the reference sequence HIV_{NL4-3} were removed with Lineup, and the sequences were concomitantly analyzed with Mfold-Phylo. The overall secondary structure is for HIV_{NL4-3}. The potential structures were then evaluated in two ways. First, Mfold-Phylo indicated nucleotide pairs (the stems of hairpins) that can form in all viruses. The boxed base-paired nucleotides are likely to form in all of these replication-competent strains. However, not all of the base pairs are identical in each of the strains. Second, p-num analysis gave an indication of the relative numbers of potential pairing partners. For this evaluation, the p-num values for the first 400 nt from the aligned virus sequences were determined. The nucleotides were then divided into three equivalent groups corresponding to those bases with the lowest, middle, and highest p-num values. p-num values ranged from 0 to 23 stable pairing partners. Low p-num values (dark lettering), intermediate values (pale lettering), and high values (circled) are indicated. Nucleotides that are struck through indicate that one or more of the viral strains do not contain a nucleotide at that particular position. Some of the strains also contain nucleotide insertions relative to HIV_{NL4-3}, and the locations and sizes of those insertions are also indicated. The *gag* start codon is underlined.

absence of hairpins 1 and 3. We think this residual encapsidation is probably due to encapsidation facilitated by additional elements. In particular, there is direct evidence that an element in the vicinity of hypothetical structure 4 or further within *gag* is necessary for encapsidation (13, 38, 48), and our data as well as those of others suggest that the region upstream of hairpin 1 also contributes to encapsidation (31, 52). Our in vivo data are also consistent with previous in vitro experiments that demonstrate cooperativity between sequences upstream and downstream of the splice donor for binding to Gag protein (11, 12, 16). Taken together, the available data reinforce the notion that the E/Ψ region is multipartite. A double-hairpin structure is important for encapsidation of murine leukemia virus (MLV) and SNV (32, 59). However, there is no apparent similarity between these double hairpins in SNV and MLV and hairpins 1 and 3 in HIV-1. HIV-1 particles do not efficiently encapsidate MLV RNA (19), and SNV particles do not efficiently encapsidate HIV-1 RNA (59).

On the basis of our results of multiple mutations that would affect the stability of hairpin 2, we conclude that this putative structure does not contribute to in vivo encapsidation. We cannot rule out the possibility that hairpin 2 functions in some other step in replication. However, virus production was not affected by mutations that would affect the stem of hairpin 2. Moreover, preliminary results using an assay to measure the efficiency of a single cycle of replication indicate that the replication efficiency of pS2 is similar to that of the wild type (data not shown).

Others have observed that deletions between the major subgenomic splice donor and the *gag* start codon have a more profound effect on encapsidation than we observed for individual mutations in the same region (4, 15, 25, 36, 38). This apparent difference in the phenotype of similar mutants likely reflects differences in the methods used to measure encapsidation efficiency. First, since there is no opportunity for infection and virus spread, our data are indicative of relative encapsidation efficiency in a single round of replication. In studies in which replication-competent virus has been examined, relatively minor differences in phenotype are likely to be amplified through multiple rounds of infection. Second, we analyzed the encapsidation of mutant RNAs in the absence of competing wild-type RNA. We expect that coexpression of wild-type RNA might magnify the mutant phenotype, since the encapsidation of nonviral RNA is enhanced in the absence of competing viral RNA (21). Third, the various protocols used to isolate viral RNA from virion particles may affect the amount of RNA isolated. Some apparent encapsidation-defective ALV mutants score as encapsidation positive when RNA is isolated in the presence of proteinase K (6). This might indicate that some mutations alter the susceptibility of encapsidated RNA to digestion with contaminating RNases. We included proteinase K treatment in our protocol for isolation of RNA from virus particles.

The retroviral genome is a dimer within the virion, but the temporal order of dimerization and encapsidation remains unclear. Dimers can be isolated from the cytoplasm of virus-producing cells, implying that dimerization may occur prior to encapsidation (17). However, some early work suggested that monomers can be isolated from wild-type virions, implying that monomers may be encapsidated and dimerize during post-budding virion maturation (14, 33). More-recent work suggests that a less stable dimer is encapsidated and becomes more stable during virus maturation (22). We did not investigate the ability of our mutants to form dimers. If dimerization is a prerequisite for encapsidation, then it is possible that some of the mutants that exhibited decreased encapsidation might be deficient in dimer formation. Interstrand quadruplex formation (8, 57), antiparallel hydrogen bonding between palindromic sequences in the loops of hairpin structures (42, 55), and interstrand, Watson-Crick base pairing between corresponding 5′ sides and 3′ sides of stems within hairpins have all been proposed as interactions that could account for dimerization. Interestingly, some of the nucleotides that constitute hairpin 1 or, alternatively, hairpin 3 have been implicated in dimer formation in vitro (18, 35, 42, 45, 47, 55). However, interstrand complementation between individual mutants with base substitutions in the 5′ and 3′ sides of stems within hairpins 1 and 3 did not increase encapsidation efficiency (Fig. 6).

Viral genomic RNA is specifically encapsidated into the assembling virus particle, and subgenomic spliced RNA is

largely excluded (38). Our results indicate that the encapsidation signal for HIV-1 spans the splice donor, so subgenomic RNAs would contain at least one element of the encapsidation signal. However, hairpin 3, which is located downstream of the splice donor, appears to play a role in discrimination between full-length and subgenomic viral RNAs. When this structure was altered by mutation, the encapsidation of spliced RNA became more efficient. We suggest that the loss of hairpin 3 in full-length viral RNA partially hampered the ability of the virus to discriminate between full-length and spliced RNAs. The data presented in this paper indicate that two particular RNA structures function in encapsidation. However, we do not yet know what is sufficient in *cis* for RNA encapsidation. The E/Ψ region appears to be multipartite, and it is altogether possible that the presence of a complete set of individual elements on full-length RNA leads to more-efficient encapsidation than the spliced viral mRNA species.

ACKNOWLEDGMENTS

We thank Katrin Talbot and Diccon Fiore for technical assistance; Kathy Boris-Lawrie, Mike Callahan, and Mike Farrell for critical reading of the manuscript; and Michael Schwartz, Dan Loeb, and Ann Palmenberg for helpful discussion.

M.S.M. was supported in part by a National Science Foundation predoctoral fellowship and a Wisconsin Alumni Research Foundation predoctoral fellowship. This work was supported by NIH grant R01 AI34733.

REFERENCES

1. **Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Wiley, A. Rabson, and M. A. Martin.** 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. **59:**284–291.
2. **Adam, M. A., and A. D. Miller.** 1988. Identification of a signal in murine retrovirus that is sufficient for packaging of nonretroviral RNA into virions. J. Virol. **62:**3802–3806.
3. **Aldovini, A., and B. D. Walker.** 1990. Techniques in HIV research. Stockton Press, New York.
4. **Aldovini, A., and R. A. Young.** 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. J. Virol. **64:**1920–1926.
5. **Armentano, D., S.-F. Yu, P. W. Kantoff, T. Von Ruden, W. F. Anderson, and E. Gilboa.** 1987. Effect of internal viral sequences on the utility of retroviral vectors. J. Virol. **61:**1647–1650.
6. **Aronoff, R., A. M. Hajjar, and M. Linial.** 1993. Avian sequences and role of the nucleocapsid Cys-His motifs. J. Virol. **67:**178–188.
7. **Aronoff, R., and M. Linial.** 1991. Specificity of retroviral RNA packaging. J. Virol. **65:**71–80.
8. **Awang, G., and D. Sen.** 1993. Mode of dimerization of HIV-1 genomic RNA. Biochemistry **32:**11453–11457.
9. **Baudin, F., R. Marquet, C. Isel, J. L. Darlix, B. Ehresmann, and C. Ehresmann.** 1993. Functional sites in the 5′ region of human immunodeficiency virus type 1 RNA form defined structural domains. J. Mol. Biol. **229:**382–397.
10. **Bender, M. A., T. D. Palmer, R. E. Gelinas, and A. D. Miller.** 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. J. Virol. **61:**1639–1646.
11. **Berkowitz, R. D., and S. P. Goff.** 1994. Analysis of binding elements in the human immunodeficiency virus type 1 genomic RNA and nucleocapsid protein. Virology **202:**233–246.
12. **Berkowitz, R. D., J. Luban, and S. P. Goff.** 1993. Specific binding of human immunodeficiency virus type 1 *gag* polyprotein and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays. J. Virol. **67:**7190–7200.
13. **Buchschacher, G. L., Jr., and A. T. Panganiban.** 1992. Human immunodeficiency virus vectors for inducible expression of foreign genes. J. Virol. **66:**2731–2739.
14. **Canaani, E., K. V. D. Helm, and P. Duesberg.** 1973. Evidence for 30-40S RNA as precursor of the 60-70S RNA of Rous sarcoma virus particles. Proc. Natl. Acad. Sci. USA **70:**401–405.
15. **Clavei, F., and J. M. Orenstein.** 1990. A mutant of human immunodeficiency virus with reduced RNA packaging and abnormal particle morphology. J. Virol. **64:**5230–5234.
16. **Clever, J., C. Sassetti, and T. G. Parslow.** 1995. RNA secondary structure and binding sites for *gag* gene products in the 5′ packaging signal of human immunodeficiency virus type 1. J. Virol. **69:**2101–2109.
17. **Darlix, J.-L., C. Gabus, and B. Allain.** 1992. Analytical study of avian reticuloendotheliosis virus dimeric RNA generated in vivo and in vitro. J. Virol. **66:**7245–7252.
18. **Darlix, J.-L., C. Gabus, M. T. Nugeyre, F. Clavel, and F. Barre-Sinoussi.** 1990. *cis*-elements and *trans*-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. J. Mol. Biol. **216:**689–699.
19. **Delwart, E. L., G. L. Buchschacher, Jr., E. O. Freed, and A. T. Panganiban.** 1992. Analysis of HIV-1 envelope mutants and pseudotyping of replication-defective HIV-1 vectors by genetic complementation. AIDS Res. Hum. Retroviruses **8:**1669–1677.
20. **Dorfman, T., J. Luban, S. P. Goff, W. A. Haseltine, and H. G. Göttlinger.** 1993. Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. J. Virol. **67:**6159–6169.
21. **Embretson, J. E., and H. M. Temin.** 1987. Lack of competition results in efficient packaging of heterologous murine retroviral RNAs and reticuloendotheliosis virus encapsidation-minus RNAs by the reticuloendotheliosis virus helper cell line. J. Virol. **61:**2675–2683.
22. **Fu, W., and A. Rein.** 1993. Maturation of dimeric viral RNA of Moloney murine leukemia virus. J. Virol. **67:**5443–5449.
23. **Harrison, G. P., E. Hunter, and A. M. L. Lever.** 1995. Secondary structure model of the Mason-Pfizer monkey virus 5′ leader sequence: identification of a structural motif common to a variety of retroviruses. J. Virol. **69:**2175–2186.
24. **Harrison, G. P., and A. M. L. Lever.** 1992. The human immunodeficiency virus type 1 packaging signal and major splice donor region have a conserved stable secondary structure. J. Virol. **66:**4144–4153.
25. **Hayashi, T., T. Shioda, Y. Iwakura, and H. Shibuta.** 1992. RNA packaging signal of human immunodeficiency virus type 1. Virology **188:**590–599.
26. **Helseth, E., M. Kowalski, D. Gabuzda, U. Olshevsky, W. Haseltine, and J. Sodroski.** 1990. Rapid complementation assays measuring replicative potential of human immunodeficiency virus type 1 envelope glycoprotein mutants. J. Virol. **64:**2415–2420.
27. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
28. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1990. Predicting optimal and suboptimal secondary structure for RNA. Methods Enzymol. **183:**281–306.
29. **Katoh, I., T. Yasunaga, and Y. Yoshinaka.** 1993. Bovine leukemia virus RNA sequences involved in dimerization and specific Gag protein binding: close relation to the packaging sites of avian, murine, and human retroviruses. J. Virol. **67:**1830–1839.
30. **Katz, R. A., R. W. Terry, and A. M. Skalka.** 1986. A conserved *cis*-acting sequence in the 5′ leader of avian sarcoma virus RNA is required for packaging. J. Virol. **59:**163–167.
31. **Kim, H.-J., K. Lee, and J. J. O'Rear.** 1994. A short sequence upstream of the 5′ major splice site is important for encapsidation of HIV-1 genomic RNA. Virology **198:**336–340.
32. **Konings, D. A. M., M. A. Nash, J. V. Maizel, and R. B. Arlinghaus.** 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.
33. **Korb, J., M. Travnicek, and J. Riman.** 1976. The oncornavirus maturation process: quantitative correlation between morphological changes and conversion of genomic virion RNA. Intervirology **7:**211–224.
34. **Koyama, T., F. Harada, and S. Kawai.** 1984. Characterization of a Rous sarcoma virus mutant defective in packaging its own genomic RNA: biochemical properties of mutant TK15 and mutant-induced transformants. J. Virol. **51:**154–162.
35. **Laughrea, M., and L. Jette.** 1994. A 19-nucleotide sequence upstream of the 5′ major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. Biochemistry **33:**13464–13474.
36. **Lever, A., H. Göttlinger, W. Haseltine, and J. Sodroski.** 1989. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J. Virol. **63:**4085–4087.
37. **Linial, M., and A. D. Miller.** 1990. Retroviral RNA packaging: sequence requirements and implications. Curr. Top. Microbiol. Immunol. **157:**125–152.
38. **Luban, J., and S. P. Goff.** 1994. Mutational analysis of *cis*-acting packaging signals in human immunodeficiency virus type 1 RNA. J. Virol. **68:**3784–3793.
39. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
40. **Mann, R., and D. Baltimore.** 1985. Varying the position of a retrovirus packaging sequence results in the encapsidation of both unspliced and spliced RNAs. J. Virol. **54:**401–407.
41. **Mann, R., R. C. Mulligan, and D. Baltimore.** 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell **33:**153–159.
42. **Marquet, R., J. C. Paillart, E. Skripkin, C. Ehresmann, and B. Ehresmann.** 1994. Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. Nucleic Acids Res. **22:**145–151.
43. **Méric, C., and S. P. Goff.** 1989. Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. J. Virol. **63:**1558–1568.

44. **Méric, C., E. Gouillard, and P.-F. Spahr.** 1988. Mutations in Rous sarcoma virus nucleocapsid protein p12(NC): deletions of Cys-His boxes. J. Virol. **62:**3328–3333.
45. **Muriaux, D., P. M. Girard, B. Bonnet-Mathoniere, and J. Paoletti.** 1995. Dimerization of HIV-Lai RNA at low ionic strength. An autocomplementary sequence in the 5′ leader region is evidenced by an antisense oligonucleotide. J. Biol. Chem. **270:**8209–8216.
46. **Page, K. A., N. R. Landau, and D. R. Littman.** 1990. Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. J. Virol. **64:**5270–5276.
47. **Paillart, J. C., R. Marquet, E. Skripkin, B. Ehresmann, and C. Ehresmann.** 1994. Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. J. Biol. Chem. **269:**27486–27493.
48. **Parolin, C., T. Dorfman, G. Palu, H. Gottlinger, and J. Sodroski.** 1994. Analysis in human immunodeficiency virus type 1 vectors of *cis*-acting sequences that affect gene transfer into human lymphocytes. J. Virol. **68:**3888–3895.
49. **Poznansky, M., A. M. L. Lever, L. Bergeron, W. Haseltine, and J. Sodroski.** 1991. Gene transfer into human lymphocytes by a defective human immunodeficiency virus type 1 vector. J. Virol. **65:**532–536.
50. **Pugatsch, T., and D. W. Stacey.** 1983. Identification of a sequence likely to be required for avian retroviral packaging. Virology **128:**505–511.
51. **Richardson, J. H., L. A. Child, and A. M. L. Lever.** 1993. Packaging of human immunodeficiency virus type 1 RNA requires *cis*-acting sequences outside the 5′ leader region. J. Virol. **67:**3997–4005.
52. **Rizvi, T. A., and A. T. Panganiban.** 1993. Simian immunodeficiency virus RNA is efficiently encapsidated by human immunodeficiency virus type 1 particles. J. Virol. **67:**2681–2688.
53. **Sakaguchi, K., N. Zambrano, E. T. Baldwin, B. A. Shapiro, J. W. Erickson, J. G. Omichinski, G. M. Clore, A. M. Gronenborn, and E. Appella.** 1993. Identification of a binding site for the human immunodeficiency virus type 1 nucleocapsid protein. Proc. Natl. Acad. Sci. USA **90:**5219–5223.
54. **Shank, P. R., and M. Linial.** 1980. Avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: characterization of a deletion in the provirus. J. Virol. **36:**450–456.
55. **Skripkin, E., J. C. Paillart, R. Marquet, B. Ehresmann, and C. Ehresmann.** 1994. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization *in vitro*. Proc. Natl. Acad. Sci. USA **91:**4945–4949.
56. **Sorge, J., W. Ricci, and S. H. Hughes.** 1983. *cis*-acting RNA packaging locus in the 115-nucleotide direct repeat of Rous sarcoma virus. J. Virol. **48:**667–675.
57. **Sundquist, W. I., and S. Heaphy.** 1993. Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA. Proc. Natl. Acad. Sci. USA **90:**3393–3397.
58. **Watanabe, S., and H. M. Temin.** 1982. Encapsidation sequences for spleen necrosis virus, and avian retrovirus, are between the 5′ long terminal repeat and the start of the *gag* gene. Proc. Natl. Acad. Sci. USA **79:**5986–5990.
59. **Yang, S., and H. M. Temin.** 1994. A double hairpin structure is necessary for the efficient encapsidation of spleen necrosis virus retroviral RNA. EMBO J. **13:**713–726.
60. **Zuker, M.** 1989. On finding all sub-optimal foldings of an RNA molecule. Science **244:**48–52.