JOURNAL OF VIROLOGY, July 1998, p. 5886–5896 Vol. 72, No. 7
0022-538X/98/$04.00+0

Functional Analysis of the Core Human Immunodeficiency Virus Type 1 Packaging Signal in a Permissive Cell Line

GEOFFREY P. HARRISON,[1]† GINO MIELE,[2] ERIC HUNTER,[1] AND ANDREW M. L. LEVER[3]*

Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294,[1] and Division of Development and Reproduction, Roslin Institute, Roslin, Midlothian, Scotland EH25 9PS,[2] and Cambridge University Department of Medicine, Addenbrooke's Hospital, Cambridge CB2 2QQ,[3] United Kingdom

Received 11 November 1997/Accepted 15 April 1998

Packaging of type C retrovirus genomic RNAs into budding virions requires a highly specific interaction between the viral Gag precursor and unique *cis*-acting packaging signals on the full-length RNA genome, allowing the selection of this RNA species from among a pool of spliced viral RNAs and similar cellular RNAs. This process is thought to involve RNA secondary and tertiary structural motifs since there is little conservation of the primary sequence of this region between retroviruses. To confirm RNA secondary structures, which we and others have predicted for this region, disruptive, compensatory, and deletion mutations were introduced into proviral constructs, which were then assayed in a permissive cell line. Disruption of either of two predicted stem-loops was found to greatly reduce RNA encapsidation and replication, whereas compensatory mutations restoring base pairing to these stem-loops had a wild-type phenotype. A GGNGR motif was identified in the loops of three hairpins in this region. Results were consistent with the hypothesis that the process of efficient RNA encapsidation is linked to dimerization. Replication and encapsidation were shown to occur at a reduced rate in the absence of the previously described kissing hairpin motif.

Retroviral 5′ untranslated leader sequences contain *cis* acting sites that are important for RNA encapsidation (1, 9, 21, 28, 35, 43, 54), dimerization (3, 12, 14, 31, 40, 60), and efficient *gag* translation (44). Retroviral RNA packaging is a highly specific process involving interactions between the viral Gag precursor and *cis* acting packaging signals in the 5′ leader sequence of their RNA genomes (5, 37, 58). The interaction is thought to involve RNA secondary structure since there is little conservation of primary sequences in the region. However, some conserved motifs have been identified. The GACG motif identified in several type C retroviruses (29) has been shown to be important for efficient encapsidation of avian spleen necrosis virus RNA (65) and murine leukemia virus RNA (45). A GAYC motif was found in the loop of a region 5′ to the *gag* initiation codon of Mason-Pfizer monkey virus and several other retroviruses (19). We (18) and others (3, 57) presented an RNA secondary-structure model for the human immunodeficiency virus type 1 (HIV-1) 5′ leader sequence region based on biochemical and enzymatic probing, comparison of the sequences of HIV-1 quasispecies, and free-energy minimization algorithms. Neither the GACG nor the GAYC motif is found in this region of HIV-1.

Electron microscopy has shown that retroviral RNAs under partially denaturing conditions are joined together in an apparent parallel orientation at a structure referred to as the dimer linkage site near the 5′ end of the genomic RNAs (4, 47). It is thought that a parallel orientation of the dimeric RNAs exists in HIV-1, and two recent publications have lent support to this theory (10, 22). Dimerization may modulate several steps of the virus life cycle, such as translation, encapsidation, recombination, and reverse transcription. Earlier work on RNA dimerization in HIV-1 using synthetic RNAs in vitro (2, 61) suggested that guanine tetrads (64) might be involved in dimer formation. This model has not been supported by in vitro studies of guanine or purine sequences in the region. Since then, stem-loop one (SL1), 5′ positions 240 to 280 (18) (now known as the kissing hairpin), has been proposed as the dimer initiation signal (7, 14, 16, 31–34, 40, 46, 49–51, 60). The present study has addressed the role of structure and sequence motifs in the HIV-1 leader in their effect on encapsidation. Putative dimer linkage sites have been included in the analysis. RNA encapsidation in nonpermissive cells (11, 38, 41, 42) has in some cases differed qualitatively and quantitatively from encapsidation in permissive cell lines (27, 34). In this study, we have examined the role of discrete RNA secondary structures in the HIV-1 core packaging signal region in encapsidation, in permissive cells, by introduction of disruptive, compensatory, and deletion mutations. RNA packaging was analyzed by using RNase protection assays (RPAs), which are both quantitative and qualitative.

MATERIALS AND METHODS

Cells and viruses. The cell line Jurkat-*tat* (56) was grown in RPMI 1640 medium supplemented with 10% fetal calf serum, penicillin, and streptomycin. COS-1 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, penicillin, and streptomycin. The infectious proviral clone HIV-1 HXB2 (13) was used in all experiments. In common with the majority of recent publications, we have numbered the viral sequence from the RNA cap site. Thus, the G of the splice donor is nucleotide (nt) 290, and the A of the *gag* initiation codon is nt 336.

Construction of mutants. The *Bgl*II fragment of pSVC21 from base 21 to base 1644 (numbering of viral sequences as described above) was cloned into the *Bam*HI site of the expression vector pBluescript KSII (Stratagene) to create a plasmid which we refer to as pKS*Bgl* II. Mutagenesis was done essentially in accordance with the method of Kunkel et al. (30). The plasmid pKS*Bgl* II was transformed into *Escherichia coli* CJ236, and a single-stranded DNA was rescued by using the manufacturer's helper phage. Synthetic oligonucleotide primers were purchased from R&D Ltd. (Oxford, United Kingdom). The sequences of mutants are shown in Fig. 1, and their locations are shown in Fig. 2.

Once the mutants had been identified by sequencing, the *Nar*I-to-*Apa*I frag-

* Corresponding author. Mailing address: Cambridge University Department of Medicine, Addenbrooke's Hospital, Hills Rd., Cambridge CB2 2QQ, United Kingdom. Phone: 44-1223-336747. Fax: 44-1223-336846. E-mail: amll1@mole.bio.cam.ac.uk.

† Present address: Ribotargets Ltd., Cambridge CB1 2JX, United Kingdom.

A1: 219 5' CCA GAG AGC TCT CTA|GAG GTG AGT ACG-3'

A2: 233 5'-CGA CGC AGG ACT C|GA GGG GCG GCG A-3'

A3: 239 5'-AGG ACT CGG CTT |AAG AGG CGA GGG G-3'

A5: 277 5'-GGG GCG GCG ACG AGT AAG TAC GCC AAA -3'

A6: 287 5'-CTG GTG AGT AGT GGT AAA ATT TT-3'

A7: 280 5'-GCG GCG ACG AGT AAG TAT ACT CAA AAT TTT-3'

A8: 306 5'-TTT TGA GAT CGG GAG GCT AGA AGG A-3'

A9: 306 5'-TTT TGA CTA GCG GAG CGA TCA AGG A-3'

A10: 306 5'-TTT TGA GAT CGG GAG CGA TCA AGG A G-3'

A11: 246 5'-GGC TTG CTG AA G CGC GCA GCC GTT CAG GCG-3'

A12: 246 5'-GGG AAC GAC AAG CGC GCA CGG CAA GAG GCG-3'

A13: 246 5'-GGG AAC GGC AAG CGC GCA GCC GTT CAG GCG-3'

ΔP1: 290 5'-GTG AGT ACG CCG GCT AGA AGG AGA GA-3'

ΔP2: 290 5'-GTG AGT ACT|GAG ATG GGT GCG AGA GC-3'

D1: 287 5'-CTG GTG AGT ACG TTT TAA ATT TT-3'

D2: 306 5'-TTT TGA CTA TTT TAG GCT AGA AGG A-3'

D3: 306 5'-TTT TGA CTA GCG GTT TTT AGA AGG A-3'

D4: 351 5'-TCA GTA TTA AGC CTT GGA GAA TTA GAT
S V L S L G E L D

FIG. 1. Nucleic acid sequences of mutants created by introducing mutations into the HIV-1 HXB2 provirus. Altered bases are underlined. The positions of deletions are indicated by vertical lines. The protein sequence of the D4 mutant is shown below the primary sequence, and the glycine-to-leucine change is underlined.

ment (from bases 186 to 1558) was excised from the mutant pKS*Bgl* II plasmids and the fragments were cloned into the unique *Nar*I-to-*Apa*I sites in the HXB2 provirus which were then revalidated by sequencing, amplifying the region from bases 202 to 464 by using a biotinylated oligodeoxynucleotide primer, 5′-B-CTG AAA GCG AAA GGG AAA CC-3′, and a nonbiotinylated oligodeoxynucleotide primer, 5′-TTCTAGCTCCCTGCTTGCCC-3′. Single-stranded DNA was prepared by bonding the amplified fragment to Dynabeads (Dynal, Cheshire, United Kingdom) and by alkali treatment using the manufacturer's recommended protocol. Single-stranded DNAs were sequenced to check that the designed mutation was present in each proviral clone.

Replication studies. Equal numbers of virus particles, as judged by using reverse transcriptase (RT), from transient COS-1 transfections were used to infect 2×10^6 Jurkat-*tat* cells for 4 h. These cells were washed in serum-free RPMI 1640 medium, and fresh RPMI 1640 medium was added (with fetal calf serum, penicillin, and streptomycin). During replication studies, two 10-μl samples of the supernatant of each culture were taken out every fourth day and frozen at −70°C until 30 days postinfection, when the RT levels were assayed simultaneously. The cells were split every fourth day, with the cell numbers kept approximately equal in all cultures.

RT levels were determined by using the RT assay of Potts (53), incorporating [α-^{32}P]dTTP (Amersham, Amersham, United Kingdom).

Protein studies. For analysis of proteins, proviral mutants were transfected into Jurkat-*tat* cells by using DEAE dextran (59). Cells were cultured in RPMI 1640 medium supplemented with 10% fetal calf serum, penicillin, and streptomycin. At 72 h posttransfection, the cells were counted and then harvested by centrifugation at 2,000 rpm for 5 min in an MSE Falcon 6/300 and resuspended in 100 μl of phosphate-buffered saline with 200 μl of sample buffer (6.25 mM Tris-HCl [pH 6.8], 2% sodium dodecyl sulfate [SDS], 10% glycerol, 5% β-mercaptoethanol, 0.02% bromophenol blue). The supernatants were filtered through 0.45-μm-pore-size filters and incubated at 4°C overnight with 0.5 volume of 30% polyethylene glycol 8000 in 0.4 M NaCl. The precipitate was collected by centrifugation at 2,000 rpm for 45 min in an MSE Falcon 6/300, and the pellet was resuspended in 100 μl of phosphate-buffered saline and 200 μl of sample buffer. Following sonication, the protein equivalent of 650,000 cells was loaded per lane (equalized quantities of supernatants were also loaded) on an SDS-polyacrylamide gel electrophoresis gel with 12.5 or 10% polyacrylamide (probing for Gag or Env, respectively). Proteins were electrophoresed for approximately 3.5 h at 50 mA and then transferred to Hybond C-extra nitrocellulose (Amersham) overnight at 250 mA. The filters were probed by using monoclonal antibodies to p55/p24 (MRCADP 313) or gp120/160 (MRCADP 323), both supplied through the United Kingdom Medical Research Council AIDS Programme. The Gag probe was used at a concentration of 3 μg/ml, and the Env probe was used at a dilution of 1 in 1,000; both were incubated for 1 h at room temperature. Bands were visualized by sheep anti-mouse horseradish peroxidase-linked whole antibody and enhanced chemiluminescence Western detection reagents (Amersham) as described in the manufacturer's instructions. Rainbow molecular weight markers (Amersham) were used.

RNA extraction and quantitation. Cytoplasmic RNAs were extracted from Jurkat-*tat* cells by rapid lysis at 4°C in NP40 buffer (50 mM Tris HCl [pH 8.00], 100 mM NaCl, 5 mM $MgCl_2$ 0.5% [vol/vol] Nonidet P-40). Cell debris and nuclei were removed by a 1-min centrifugation at 13,000 rpm in a microcentrifuge. The supernatants were adjusted to 0.2% SDS and 125 μg of proteinase K per ml, incubated at 37°C for 15 min, and extracted twice with phenol-chloroform and once with chloroform. Virion RNAs were prepared from the supernatants of chronically infected Jurkat-*tat* cells. RT assays were carried out on two 10-μl aliquots from the supernatant of each culture. Based on these data, equivalent numbers of particles were collected from each supernatant (about 2 ml from the

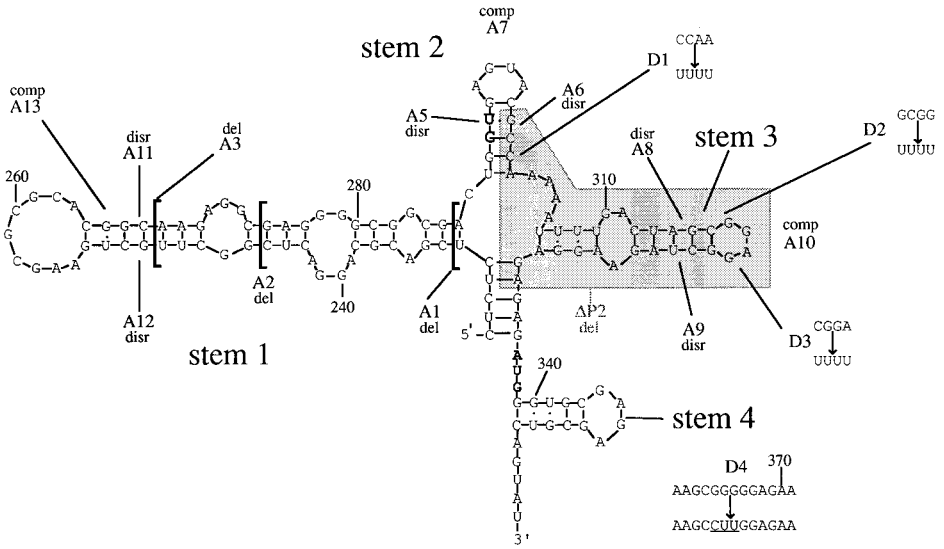
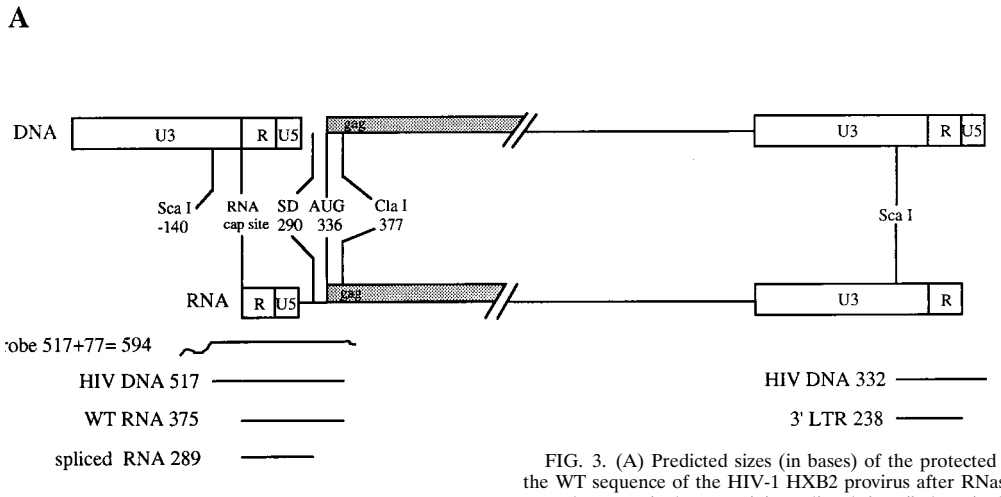


FIG. 2. Positions of mutations within the RNA secondary structures of the HIV-1 HXB2 5′ leader sequence showing our original prediction for the structure of SL2 (stem 2), which differs from that of Sakaguchi et al. (57). The sequence deleted in the ΔP2 mutation is shaded. Abbreviations: del, deletion: disr, disruption: comp, compensatory mutation. Stems 1 to 4, SL1 to SL4.



B

A1	230	A11	262
A2	243	A12	246
A3	248	A13	246
A5	290	ΔP1	299
A6	295	ΔP2	296
A7	290	D1	297
A8	310	D2	313
A9	319	D3	320
A10	310	D4	361

FIG. 3. (A) Predicted sizes (in bases) of the protected RNA fragments for the WT sequence of the HIV-1 HXB2 provirus after RNase protection assays. LTR, long terminal repeat. (B) Predicted sizes (in bases) of the protected RNA fragments which characterize mutant sequences after RNase protection assays using the WT riboprobe.

wild-type [WT] virus supernatant). HIV-1 particles were harvested as described above and then resuspended in 0.5 ml of TNE (10 mM Tris HCl, 150 mM NaCl, 1 mM EDTA [pH 7.5]). This suspension was layered over an equal volume of TNE containing 20% sucrose and centrifuged at 98,000 × *g* in a Beckman TLA45 rotor for 2 h. The pellets were resuspended in TNE and stored at 4°C while another RT assay was carried out on the suspension. Equivalent amounts of particles were then used for RNA extractions. The suspensions were made up to equal volumes with TNE; an equal volume of 2X proteinase K buffer (100 mM Tris HCl [pH 7.5], 200 mM NaCl, 20 mM EDTA, 2% SDS, 200 μg of proteinase K per ml, 200 μg of tRNA per ml) was then added and the mixture was incubated for 30 min at 37°C. Two phenol-chloroform extractions were carried out. RNAs from both cytoplasmic and virion samples were precipitated under ethanol and stored at −70°C.

RPAs were carried out with the RPA II kit supplied by Ambion (Austin, Tex.). For expression of the HIV-1 antisense riboprobe, the plasmid pKSΨCS (26), which has the *Sca*I-to-*Cla*I fragment of pSVC21 (from 140 bases upstream of the RNA cap site to 377 bases downstream) cloned into the *Eco*RV and *Cla*I sites of pBluescript KSII, was used. The plasmid was linearized with *Xba*I, and RNA was transcribed from the T3 promoter by using an in vitro transcription kit (Promega) incorporating Redivue [α-^{32}P]UTP (Amersham). The riboprobe contained 517 bases of WT HIV-1 sequence, with 77 bases from the vector making a total length of 594 nt (Fig. 3A).

A riboprobe for the human actin mRNA, pTRI-B actin-125-human (Ambion), was used to determine the levels of actin mRNA in cytoplasmic RNA preparations and to determine whether any actin mRNA was encapsidated. The actin probe was 160 nt in length and protected RNAs of 127 nt. The HIV-1 riboprobe distinguished between spliced RNA (289 bases), unspliced RNA (WT, 375 bases), and 3′ long-terminal-repeat RNA (236 bases) (Fig. 3A). During the RPAs, the WT riboprobe was digested at the sites of mutations, resulting in protected fragments of characteristic length. The predicted sizes of the larger protected unspliced RNA fragments are detailed in Fig. 3B.

RNA markers were transcribed from the Century marker template (Ambion) with T7 polymerase incorporating Redivue [α-^{32}P]UTP (Amersham). The intensities of protected HIV-1 RNAs on gels were determined with a laser densitometer (Molecular Dynamics model 300A) and were calculated as a percentage of that of WT. The results of at least three separate experiments were used to calculate the standard deviations.

Secondary-structure predictions. In this paper, we refer to the stem-loop from 232U to 286A as SL1, structures in the region from 288U to 301A as SL2, the stem-loop from 306U to 330A as SL3, and the stem-loop from 339G to 352C as SL4. The nucleic acid folding program MFold (24, 66) adapted for GCG (Genetics Computer Group, University of Wisconsin, Madison) was used for free-energy minimization predictions with the graphical presentation of Squiggles (48) in the GCG program Plotfold.

RESULTS

Figures 4 to 6 show Western blots of viral proteins produced by the mutants. None of the mutations abolish protein production as judged by analysis of cells and supernatants from transient transfections. Lower-than-WT levels of Gag and Env were detected on Western blots for mutants A11 (SL1) and A12 (SL1) as well as A8 (SL3) and ΔP1 (SL3). D4, which disrupts Gag coding, led to a decrease in Gag but not Env protein that was detectable in both cells and supernatants. The

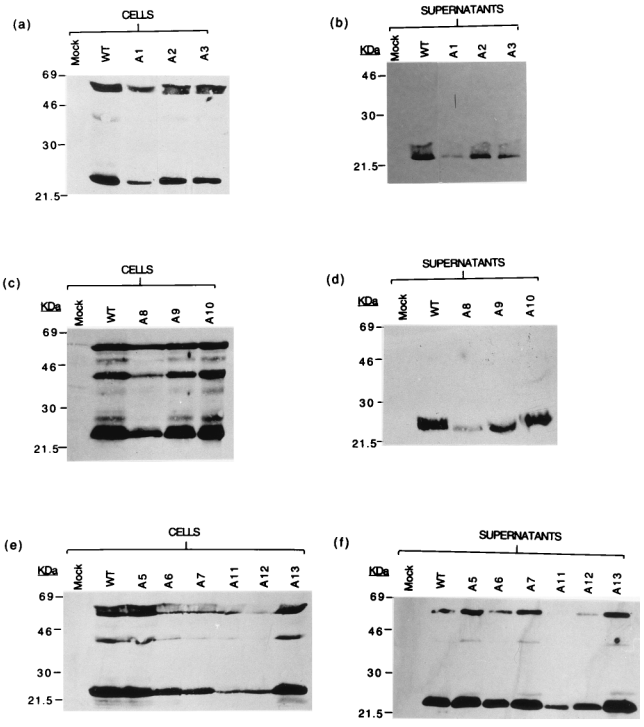


FIG. 4. Western blots of proteins extracted from cytoplasms (a, c, and e) and supernatants (b, d, and f) of Jurkat-*tat* cells transfected with the WT and mutant HIV-1 proviruses A1 to A3 and A5 to A13, probed with a monoclonal antibody to p55/24. Mock, nontransfected cells.

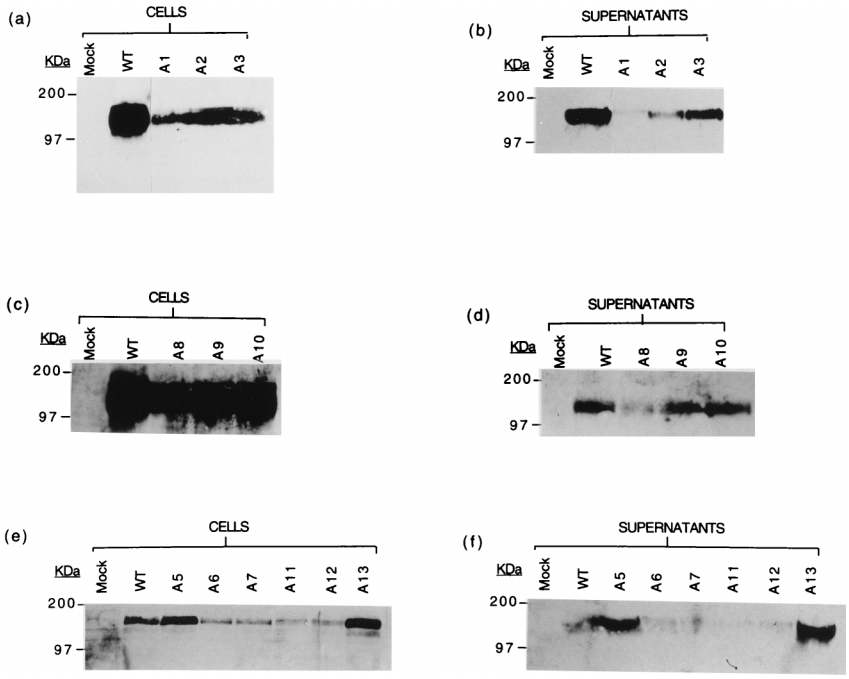


FIG. 5. Western blots of proteins extracted from the cytoplasms (a, c, and e) and supernatants (b, d, and f) of Jurkat-*tat* cells, transfected with the WT and mutant HIV-1 proviruses A1 to A3 and A5 to A13, probed with a monoclonal antibody to gp120/160. Mock, nontransfected cells.

integrity of SL1 and SL3 thus has an influence on viral protein production.

The A1 mutant was lethal, and no functional infective viruses were produced. Other than this mutant, sufficient viral particles were produced from infected Jurkat-*tat* cells to analyze viral encapsidation in equivalent (RT-equalized) quantities of viral particles. Viral replication assays are shown in Fig. 7. Replication was monitored for up to 30 days. WT virus replication showed a peak value at about day 16. RPAs of cellular and viral RNA are shown in Fig. 8 and 9. These assays distinguish spliced and unspliced species. The relative efficiencies of viral RNA packaging are shown in Fig. 10.

Replication and encapsidation closely mirrored each other.

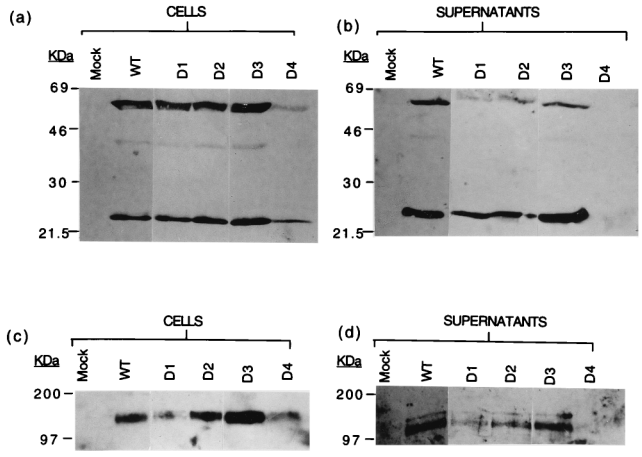


FIG. 6. Western blots of proteins extracted from the cytoplasms (a and c) and supernatants (b and d) of Jurkat-*tat* cells transfected with the WT and mutant HIV-1 proviruses D1 to D4, probed with a monoclonal antibody to p55/24 (a and b) or to gp120/160 (c and d). Mock, nontransfected cells.

A2 (SL1) and A11 (SL1) severely impaired replication and encapsidation, the less severe truncation of SL1 (A3) having a less marked effect, and A13, the compensatory mutation for A11, restored WT levels of replication and encapsidation. This is likely due to the reformation of the base pairing in SL1, albeit with an altered sequence. A8 (SL3) and ΔP2 (SL3) both profoundly reduced encapsidation and packaging, and A10, the compensatory mutant of A8, restored WT levels of both. None of the purine motif mutants D1 to D4 significantly affected viral replication or RNA encapsidation. Disruption of SL2 (A5 and A6) had little effect, but combined mutations of both sides of the helix, which maintained structure but altered sequence (A7), impaired RNA encapsidation and viral replication.

DISCUSSION

This study has documented the effects of a large number of mutations in the packaging signal region of HIV-1 on encapsidation and revealed the functional importance of SL1 and SL3. Some of these mutations also had effects on viral protein and particle production. This might be secondary to the effects on RNA stability, or, in some cases, there may have been minor effects on splicing efficiency. Such effects may contribute to the rate of viral spread in a culture. However, all of the encapsidation data presented have been obtained from equivalent quantities of viral particles; thus, clear conclusions on the effects of the mutations on encapsidation alone and on the role of the particular structures within this region on retroviral RNA packaging can be drawn.

SL1. The functionally important part of SL1 in HXB2 probably lies between bases 243A and 277G (Fig. 2) even though the stem-loop has the potential to form three more GC base pairs. This is because the equivalent structure in chimpanzee immunodeficiency virus (CIV) (23) and the clade O isolates ANT70C and MVP5180 (15, 17) (Fig. 11) starts at base 243A

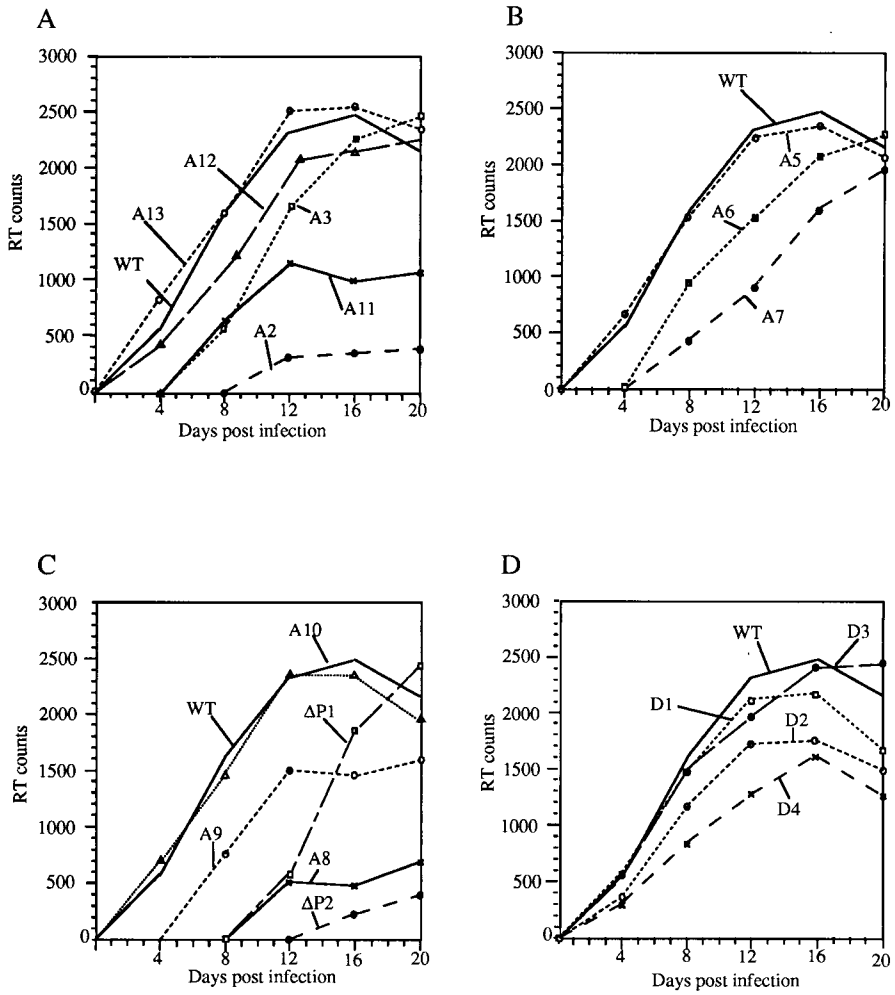


FIG. 7. RT activity in the supernatants of Jurkat-*tat* cells infected with the WT and mutant proviruses (counts per second of [^{35}S]TTP incorporation per 10 μl) plotted against days postinfection. (A) Mutations to SL1; (B) mutations to SL2; (C) mutations to SL3; (D) mutations to purine motifs.

in HXB2 and ends at base 277G (Fig. 2). Kim et al. (28) deleted bases 216 to 278 in the isolate HIV-1_{LAI} and found that replication was delayed by 4 days in MT4 cells. Paillart et al. (50) deleted bases 243 to 277 and also bases 248 to 270 and found low levels of infectivity (the pattern of replication was not reported).

The kissing hairpin is thought to be the signal for initiation of dimerization of genomic RNAs (7, 14, 16, 31, 32, 34, 40, 49–51, 60). Our deletion mutants A2 and A3 both maintained palindromes at the point where bases were removed; therefore, it could be argued that these novel palindromes substituted for the absence of the WT palindrome. However, the deletion mutants Δ248–270 and Δ243–277 of Paillart et al. (50) did not contain novel palindromes yet retained some ability to infect permissive cells, albeit at a level much lower than that of WT. This suggests that in the absence of the palindrome, an alternative method of initiation of dimerization occurs or that monomer RNA can be encapsidated. The fact that alterations to palindromic sequences in the loop of the kissing hairpin decrease but do not abolish replication (7, 10, 16, 50) is consistent with the suggestion that the kissing hairpin is not an absolute requirement for either dimerization or encapsidation (33). It has been shown that sequences upstream of SL1 contribute to dimerization (7) and encapsidation (28, 42). Previous studies have also shown that retroviral dimers are linked at many locations within the 5′ end of the genomic RNA (39). The processes of encapsidation and dimerization may be tightly linked (14, 34). Such a tight linkage would explain why sequences upstream of the splice donor (SD) sequence are required for encapsidation of genomic RNAs since there are dimerization signals upstream of the SD. The WT phenotype of the compensatory mutant A13 shows that the structure of the stem of the kissing hairpin rather than its sequence is critically important. McBride and Panganiban (41) made a disruption mutation in SL1 that was slightly different from that in our A12 mutant, changing 248C to 248A, 249U to 249A, and 251G to 251U. This disruption decreased RNA packaging in HeLa cells by about one-fifth compared to that of WT. The difference between our disruption mutants A11 and A12 may be due to the effect of these disruptions on RNA secondary structure. Computer predictions (24, 66) for the secondary structure of the disruption mutants A11 and A12 did not provide any explanation of these differences in terms of the maintenance of palindrome regions (data not shown). We note that mutations to the purine-rich bulge of SL1 (245G to 245U, 268A to 268U, 269G to 269U, and 270G to 270U) were shown to affect RNA packaging in nonpermissive cells (11).

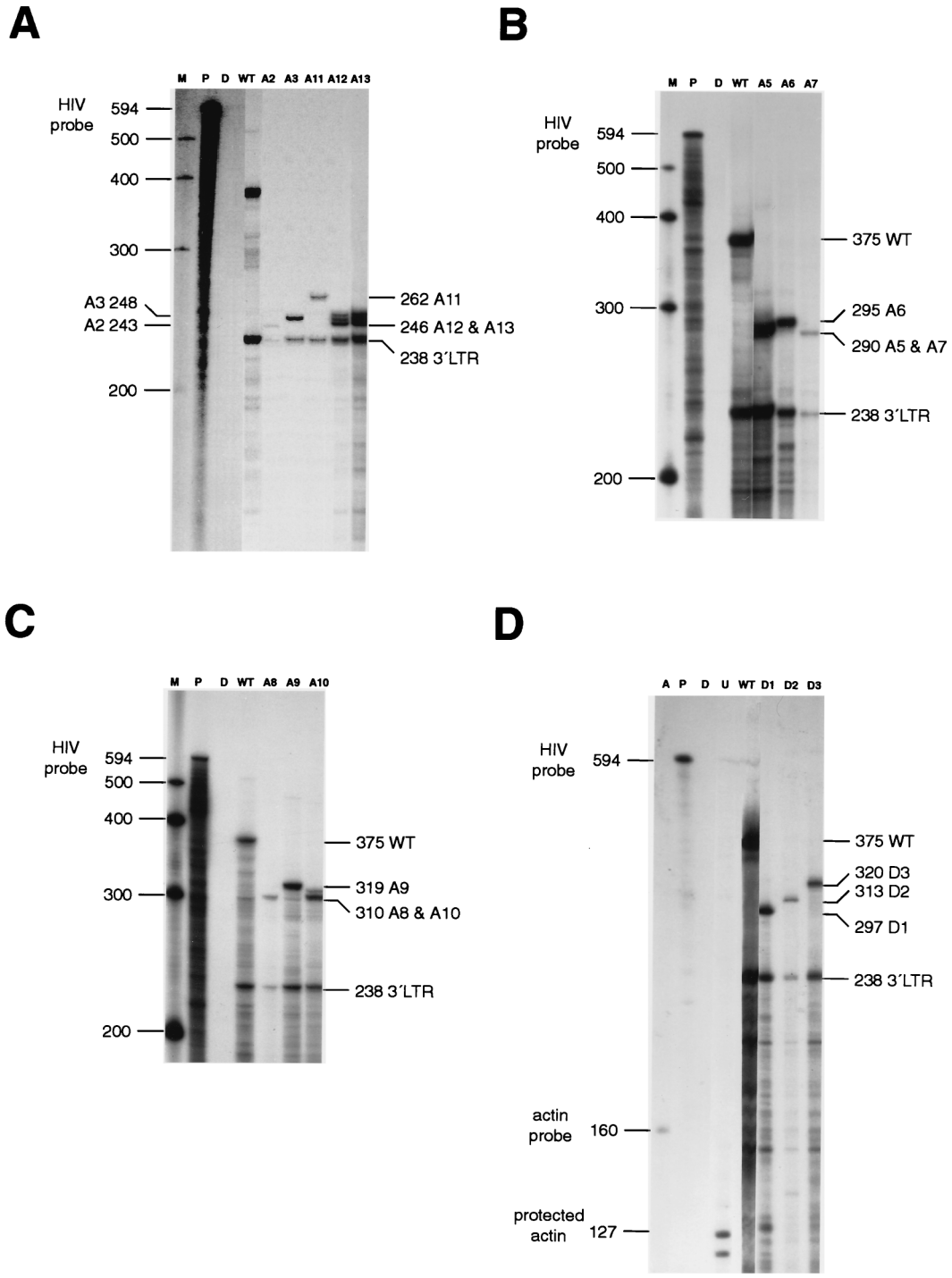


FIG. 8. RPAs of RNAs extracted from virions in the supernatants of chronically infected Jurkat-*tat* cells. Abbreviations: M, RNA size markers; P, undigested riboprobe; D, digested riboprobe; U, uninfected cells; LTR, long terminal repeat. Lanes containing relevant mutants are indicated.

SL2. This stem-loop, which contains the major splice donor signal, has two possible structures: one is shown in Fig. 2, and the other, first documented by Sakaguchi et al. (57), is shown in Fig. 12 (see below). The alternative structure for this stem-loop places the SD at the end of a stem-loop. In this prediction, there is a stable stem with a loop containing the motif GGUGA. We were interested in addressing the question of whether either of these two possible structures has functional importance. Our disruption mutants (A5 and A6) and the compensatory mutations in this stem-loop were based on the structure shown in Fig. 2. Sequence comparisons between the HXB2 sequence and divergent HIV sequences show that the only part of this region which is conserved between HIV HXB2 (Fig. 2) and CIV (23) and the clade O isolates (ANT70C and MVP5180) (15, 17) is the sequence GGUGAG containing the SD signal (Fig. 11); even so, our mutant A5 altered the se-

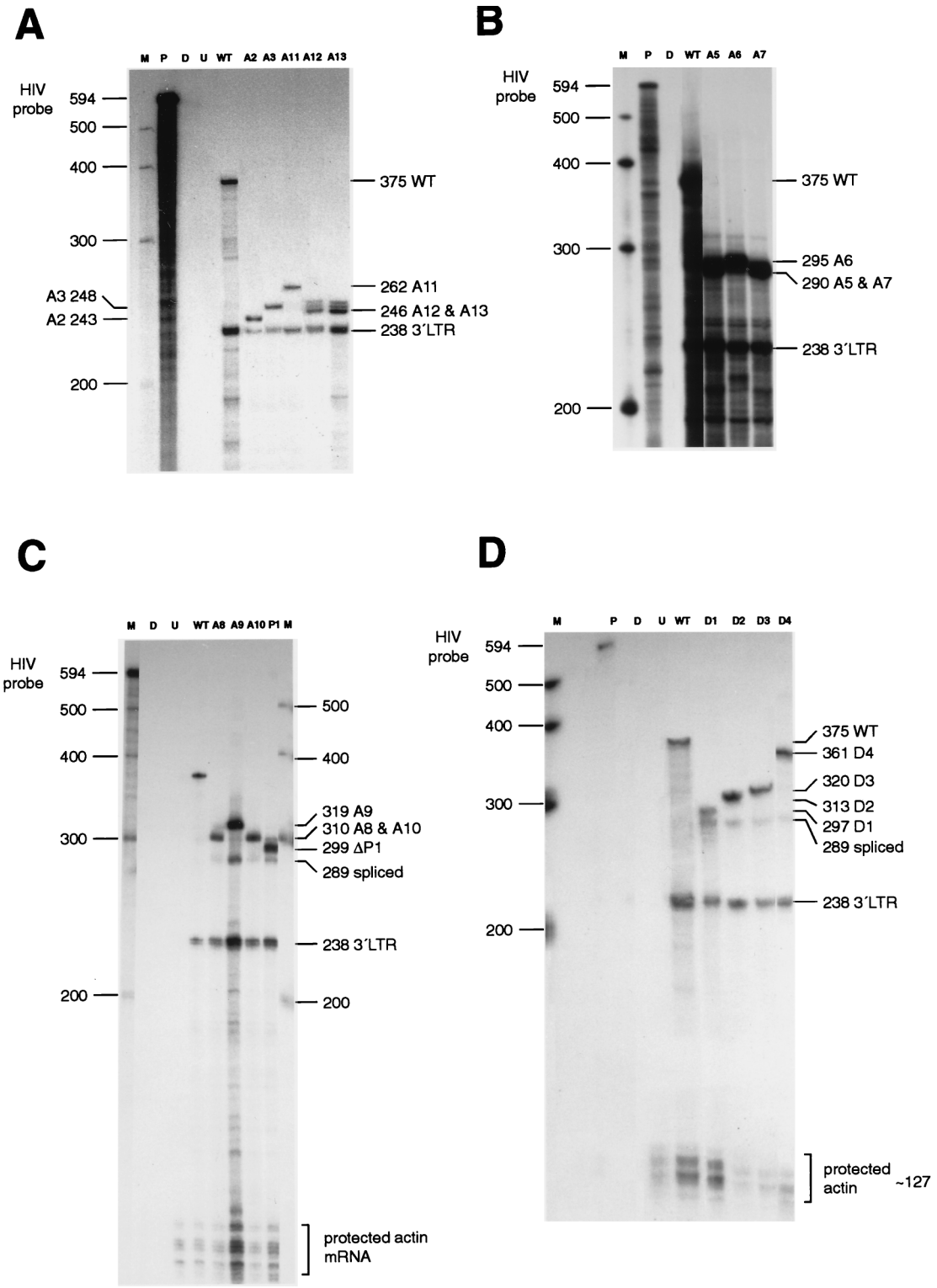


FIG. 9. RPAs of RNAs extracted from the cytoplasms of chronically infected Jurkat-*tat* cells. Abbreviations: M, RNA size markers; P, undigested riboprobe; D, digested riboprobe; U, uninfected cells; LTR, long terminal repeat. Lanes containing relevant mutants are indicated.

quence motif from GGUGAG to AGUAAG with little effect on replication. The compensatory mutant in this stem (A7) (Fig. 2) did not have a WT phenotype; therefore, either our original suggested structure of this region is not functionally important or the primary sequence is more important than its secondary structure. Kim et al. (28) deleted bases 293A to 302A and found that replication of this virus was delayed slightly. RT activity from the mutant provirus in MT4 cells

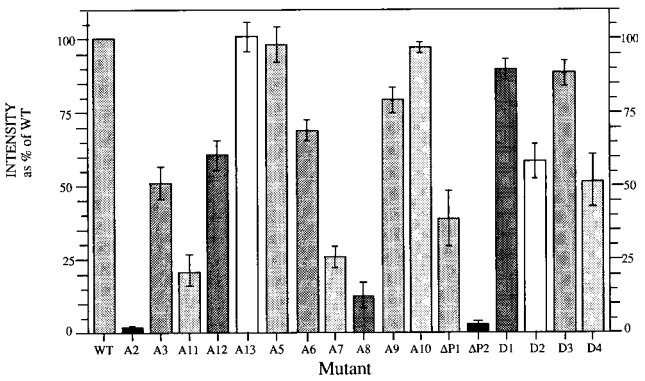


FIG. 10. Intensity of protected RNA bands, after RPAs with virion RNAs, shown as a percentage of the intensity of the band from WT RNA.

peaked at 9 days as opposed to 6 days for the WT. The predicted effects of mutants A5, A6, A7, and D1 on the Sakaguchi structure are shown in Fig. 12. Mutants A6 and A7 disrupt the Sakaguchi model for this region, whereas the A5 and D1 mutants still present the GU splice donor motif on the end of a stem-loop. This strongly suggests that the Sakaguchi prediction is the structure which has functional importance. We note that our mutant A1 also disrupts the Sakaguchi prediction for SL2. The Sakaguchi model of this region is conserved in CIV (Fig. 12), whereas our original prediction (18) is not. Neither structure is conserved in the equivalent region of the clade O isolates, in which the SD is predicted to be 3′ to a purine loop. The effect of sequence variations in the naturally occurring HIV-1 isolates RF and Z2Z6 is shown in Fig. 12. The lack of conservation of this sequence and structure in this region suggests that SL2 has less functional importance than SL1 and SL3.

SL3. Common to all of the secondary-structure predictions to date (3, 18, 57) is a stem-loop structure including bases 312C to 325G. The A at position 319 changes to U in CIV and the clade O isolates; there is also a compensatory mutation in the clade O viruses where bases 314A and 323U are changed to G and C, respectively. The loop of SL3 contains the conserved motif GGNGR. In CIV and the clade O HIV isolates, there is a third such motif, GGCGR, in the extra stem-loop that is found in these viruses (Fig. 11). This suggests that the

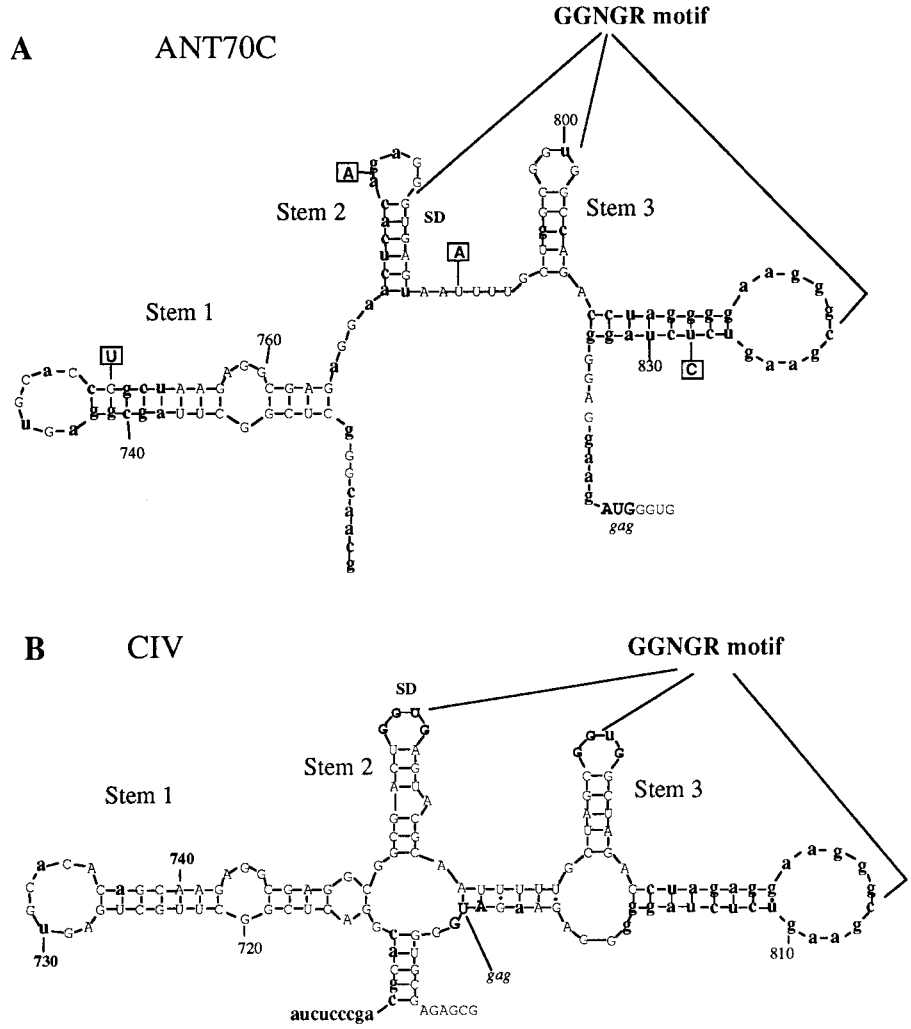


FIG. 11. RNA secondary-structure predictions done by use of the GCG program MFold. Bases which differ from those of HIV-1 HXB2 are in lowercase letters. (a) The divergent HIV-1 isolate ANT70C. Bases in MVP5180 which differ from those of ANT70C are shown in boxes. (b) CIV.

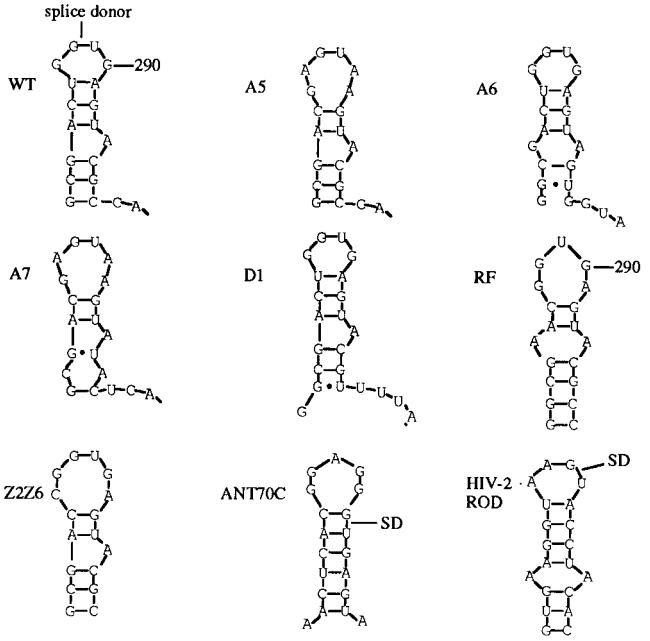


FIG. 12. Predictions of RNA secondary structures within SL2 of RNAs from mutant proviruses, based on the structure for this region first documented by Sakaguchi et al. (57), by using the output program MFold. Predictions of similar structures in the equivalent regions of the HIV-1 isolates RF, Z2Z6, ANT70C, and HIV-2 ROD are shown.

GGNGR motif is important in primate lentiviruses either in the encapsidation process or in another essential part of the replication cycle. It is possible that the presence of just one GGNGR motif is sufficient for the normal life cycle to occur. The ΔP1 deletion (bases 298A to 319A inclusive) first demonstrated that this region is required for efficient packaging (35). A mutant with a larger (298G to 332A) deletion (ΔP2) replicated detectably after only 20 days in Jurkat-*tat* cells, similar to the A2 deletion mutant. Deletions in the same area assayed by other groups are not directly comparable, but they do indicate that this region is important for HIV-1 RNA encapsidation. Clavel and Orenstein (9) deleted bases 303A to 332G from the sequence of isolate HIV_{NL4-3} and found that while WT levels of particles were produced in SL480 cells by this mutant, these particles could not infect A3.01 CD4-positive lymphoid cells. Kim et al. (28) deleted bases 308U to 328G from the sequence of HIV_{LAI} and found that replication was delayed by 5 days compared to that of WT in MT4 cells due to a packaging defect, as detected by slot blot analysis. A very similar deletion mutant, mutant pA4HXB, of Aldovini and Young (1) (with bases 295U to 315G deleted from isolate HXB2 sequence) did not replicate in H9 cells, whereas the present results show that replication of the deletion mutant ΔP1 (with bases 298A to 319A inclusive deleted) was delayed only in Jurkat-*tat* cells. This suggests that there are significant differences in the phenotypes of mutants when assayed in these two permissive lines, H9 and Jurkat-*tat*. Similarly, pA3HXB, the 39-bp deletion mutant of Aldovini and Young (from 295U to 333G deleted) did not replicate, whereas our ΔP2 deletion mutant (from 298G to 332A deleted) did so after a long delay. However, we note that the mutants were not identical.

Free-energy minimization calculations for the disruption mutant A8 predicted no stable secondary structure for the region between nt 310 and nt 320 in mutant A8 (data not shown), whereas the disruption mutant A9 could form a weak structure that presents the GGAGC motif at the end (data not shown). Perhaps the difference between the packaging efficiency and replication rates of disruption mutants A8 and A9 are due to the fact that A8 cannot present the GGAGG motif on the end of a stem whereas A9 can form a weak secondary structure which presents the motif GGAGC in a loop (data not shown). The compensatory mutant A10 showed that the structure of this stem is sufficient to restore the WT phenotype independent of its sequence. Sequences downstream of the *gag* initiation codon have been reported to be required for efficient RNA encapsidation (38), but apart from our D4 mutant, we have not investigated this region.

Purine-rich sequences. Awang and Sen (2), Sundquist and Heaphy (61), and Marquet et al. (40) proposed that dimerization occurs as a result of highly stable purine-purine interactions. This work was carried out on synthetic RNAs in vitro in the absence of proteins and may not be directly relevant in vivo. Our proviral constructs D1 to D4 abolished four purine-rich sequences identified by Sundquist and Heaphy (61) but did not produce effects on replication that were consistent with interference with dimerization. A guanine-to-uridine change at base 365 introduced by Haddrick et al. (16), which interrupted a run of five G residues, did not have a deleterious effect on replication of isolate HIV_{NL4-3}. Our alteration of the same run of five G's to CUUGG (D4) had the greatest effect of all our purine motif mutants; however, a significant contributor to this effect may be the coding change of a glycine to a leucine in *gag*. Removal of a purine bulge in SL1 of HIV-1 was found to affect RNA encapsidation in a nonpermissive cell line (11). We conclude that the removal of any single G-rich purine motif in this region has little effect on encapsidation or replication. Similarly, dimer linkage in bovine leukemia virus and HIV-2 appears to be independent of purine motifs (6, 25). However, predicted loops and bulges within RNA structures of the primate immunodeficiency lentiviruses are very rich in purines (18a) and purines have previously been shown to be important for highly specific binding of protein to RNA in other systems. These include the HIV-1 Rev binding to a purine-rich bubble of the *rev*-responsive element (20, 52), yeast ribosomal binding protein L32 recognition of the 5′ end of its mRNA (36, 63), and recognition of an RNA hairpin by the MS2/R17 coat protein (55, 62). It seems likely that Gag may interact with unpaired purines such as GGNGR motifs during encapsidation.

We conclude that the stem-loop structures at the ends of SL1 and SL3 are both important for efficient encapsidation in permissive cells independent of their stem sequences. Our data support and augment those of McBride and Panganiban (41) for HeLa cells and of Clever and Parslow for human osteosarcoma cells and 293T cells (11); moreover, we have formally excluded the possibility of promiscuous RNA encapsidation, by analyzing encapsidation of actin mRNA in this permissive cell line. In contrast to McBride and Panganiban, we found that compensatory mutations to these stem-loops completely restored the WT phenotype. Like them, we found that in the absence of either SL1 or SL3, some RNA encapsidation still occurs and is highly specific, implying that the RNA encapsidation signal is multipartite and diffuse (8, 11, 41). We did not investigate the ability of our mutants to form dimers. Our data show that deletions to the kissing hairpin (SL1) affect RNA encapsidation. Other workers have shown that alterations to SL1 have an effect on dimerization (7, 14, 16, 31–34, 40, 46, 49–51, 60). It appears that mutations to SL1 affect both RNA dimerization and RNA encapsidation in the same way, which suggests that the processes of dimerization and RNA encapsidation are linked. A GGNGR motif is found in the loops of

SL2 and SL3 and in an additional stem-loop found between the SD and the *gag* initiation codon in CIV and the known clade O isolates. Disruption of one copy of this motif has little effect on virus replication and RNA encapsidation, suggesting that multiple copies of the motif may confer a degree of functional redundancy or that they have a role other than encapsidation in the virus life cycle.

These data have led us to reevaluate structural predictions made by ourselves and others. The structure which we previously referred to as stem-loop 1 (18), in which bases 227C to 231C were predicted to be bound to bases 335G to 331G, is poorly conserved in CIV and is not conserved at all in the known clade O isolates (Fig. 11); therefore, it is probably not functionally important in HXB2. The structure labeled SL4 (stem 4) in Fig. 2 (38) is not conserved in CIV or the clade O isolates and therefore is probably not functionally important.

ACKNOWLEDGMENTS

We thank Teresa Barnes for secretarial work and J. Greatorex, S. Hassard, M. Sakalian, S. Heaphy, and W. Sundquist for helpful discussions.

This work was supported by the Medical Research Council (United Kingdom), The Sykes Trust, and EC Biomed grant BMH4-CT96-0675.

REFERENCES

1. **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.
2. **Awang, G., and D. Sen.** 1993. Mode of dimerization of HIV-1 genomic RNA. Biochemistry **32:**11453–11457.
3. **Baudin, F., R. Marquet, C. Isel, J.-L. Darlix, B. Ehresmann, and C. Ehresmann.** 1993. Functional sites in the 5′ region of HIV-1 RNA form defined structural domains. J. Mol. Biol. **229:**382–397.
4. **Bender, W., and N. Davidson.** 1976. Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type C type oncornavirus RNA molecules. Cell **7:**595–607.
5. **Berkhout, B.** 1997. Structure and function of the HIV leader RNA. Prog. Nucleic Acid Res. **54:**1–34.
6. **Berkhout, B., B. B. Essink, and I. Schoneveld.** 1993. *In vitro* dimerisation of HIV-2 leader RNA in the absence of PuGGAPuA motifs. FASEB J. **7:**181–187.
7. **Berkhout, B., and J. L. B. van Wamel.** 1996. Role of the DIS hairpin in replication of human immunodeficiency virus type 1. J. Virol. **70:**6723–6732.
8. **Berkowitz, R. D., Å. Öhagen, S. Höglund, and S. P. Goff.** 1995. Retroviral nucleocapsid domains mediate the specific recognition of genomic viral RNAs by chimeric *Gag* polyproteins during RNA packaging in vivo. J. Virol. **69:**6445–6456.
9. **Clavel, 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.
10. **Clever, J. L., M. L. Wong, and T. G. Parslow.** 1996. Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus type 1 RNA. J. Virol. **70:**5902–5908.
11. **Clever, J. L., and T. G. Parslow.** 1997. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. J. Virol. **71:**3407–3414.
12. **Darlix, J.-L., C. Gabus, M. T. Nugeyre, F. Clavel, and F. Barré-Sinoussi.** 1990. *cis*-elements and *trans*-acting factors involved in the RNA dimerisation of HIV-1. J. Mol. Biol. **216:**689–699.
13. **Fisher, A. G., E. Collati, L. Ratner, R. C. Gallo, and F. Wong-Staal.** 1985. A molecular clone of HTLV III with biological activity. Nature (London) **316:**262–265.
14. **Fu, W., R. J. Gorelick, and A. Rein.** 1994. Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. J. Virol. **68:**5013–5018.
15. **Gürtler, L. G., P. H. Hauser, J. Eberle, A. von Brunn, S. Knapp, L. Zekeng, J. M. Tsague, and L. Kaptue.** 1994. A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. J. Virol. **68:**1581–1585.
16. **Haddrick, M., A. L. Lear, A. J. Cann, and S. Heaphy.** 1996. Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. J. Mol. Biol. **259:**58–68.
17. **Haesevelde, M. V., J. L. Decourt, R. J. DeLeys, B. Vanderborght, G. der Groen, H. van Neuverswiyn, and E. Saman.** 1994. Genomic cloning and complete sequence analysis of a highly divergent African human immunodeficiency virus isolate. J. Virol. **68:**1586–1596.
18. **Harrison, G. P., and A. M. L. Lever.** 1992. The 5′ packaging signal and major splice donor region of human immunodeficiency virus type 1 have a conserved stable secondary structure. J. Virol. **66:**4144–4153.

18a. **Harrison, G. P., and A. M. L. Lever.** Unpublished observations.

19. **Harrison, G. P., E. Hunter, and A. M. L. Lever.** 1995. Secondary structure 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.
20. **Heaphy, S., J. T. Finch, M. Gait, J. Karn, and M. Singh.** 1991. HIV-1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. Proc. Natl. Acad. Sci. USA **88:**7366–7370.
21. **Hayashi, T., Y. Ueno, and T. Okamoto.** 1993. RNA packaging signal of HIV-1. FEBS Lett. **327:**213–218.
22. **Höglund, S., Å. Öhagen, J. Goncalves, A. T. Panganiban, and D. Gabuzda.** 1997. Ultrastructure of HIV-1 genomic RNA. Virology **233:**271–279.
23. **Huet, T., R. Cheynier, A. Meyerhans, G. Roelants, and S. Wain-Hobson.** 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature (London) **345:**356–359.
24. **Jaeger, J. A., D. H. Turner, and M. Zucker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
25. **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.
26. **Kaye, J. F., J. H. Richardson, and A. M. L. Lever.** 1995. *cis*-acting sequences involved in human immunodeficiency virus type 1 RNA packaging. J. Virol. **69:**6588–6592.
27. **Kaye, J. F., and A. M. L. Lever.** 1996. *trans*-acting proteins involved in RNA encapsidation and viral assembly in human immunodeficiency virus type 1. J. Virol. **70:**880–886.
28. **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.
29. **Konings, D. A., M. A. Nash, J. V. Maizel, and R. B. Arlinghaus.** 1992. Novel GACG-hairpin motif in the 5′ untranslated region of type C retroviruses related to murine leukemia virus. J. Virol. **66:**632–640.
30. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**367–382.
31. **Laughrea, M., and L. Jetté.** 1994. A 19 nucleotide sequence upstream of the major splice donor is part of the dimerization domain of HIV-1 genomic RNA. Biochemistry **33:**13464–13474.
32. **Laughrea, M., and L. Jetté.** 1996. HIV-1 genomic dimerisation: formation kinetics and thermal stability of dimeric HIV-1 LAI RNAs are not improved by the 1–232 and 292–790 regions flanking the kissing loop domain. Biochemistry **35:**9366–9374.
33. **Laughrea, M., and L. Jetté.** 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 dispensible for dimer formation. Biochemistry **35:**1589–1598.
34. **Laughrea, M., L. Jetté, J. Mak, L. Kleiman, C. Liang, and M. A. Wainberg.** 1997. Mutations in the kissing loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA dimerization. J. Virol. **71:**3397–3406.
35. **Lever, A. M. L., H. Gottlinger, 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.
36. **Li, H., and S. A. White.** 1997. RNA aptamers for yeast ribosomal protein L32 have a conserved purine-rich internal loop. RNA **3:**245–254.
37. **Linial, M. L., and A. D. Miller.** 1990. Retroviral RNA packaging: sequence requirements and implications, p. 125–152. *In* R. Swanstrom and P. K. Vogt (ed.), Retroviruses: strategies of replication. Springer-Verlag, Berlin, Germany.
38. **Luban, J., and S. P. Goff.** 1994. Mutational analysis of *cis* acting packaging signals in HIV-1 RNA. J. Virol. **68:**3784–3793.
39. **Mangel, W. F., H. Delius, and P. H. Duesburg.** 1994. Structure and molecular weight of the 60-70S RNA and the 30-40S RNA of Rous sarcoma virus. Proc. Natl. Acad. Sci. USA **71:**4541–4545.
40. **Marquet, R., J.-C. Paillart, E. Skripkin, C. Ehresmann, and B. Ehresmann.** 1994. Dimerization of HIV-1 RNA involves sequences located upstream of the splice donor site. Nucleic Acids Res. **22:**145–151.
41. **McBride, M. S., and A. T. Panganiban.** 1996. The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. J. Virol. **70:**2963–2973.
42. **McBride, M. S., M. D. Schwartz, and A. T. Panganiban.** 1997. Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of *cis* elements required for encapsidation. J. Virol. **71:**4544–4554.
43. **McCann, E., and A. M. L. Lever.** 1997. Location of *cis*-acting signals important for RNA encapsidation in the leader sequence of human immunodeficiency virus type 2. J. Virol. **71:**4133–4137.
44. **Miele, G., A. Mouland, G. P. Harrison, E. Cohen, and A. M. L. Lever.** 1996. The human immunodeficiency virus type 1 5′ packaging signal structure

affects translation but does not function as an internal ribosome entry site structure. J. Virol. **70:**944–951.
45. **Mougel, M., and E. Barklis.** 1997. A role for two hairpin structures as a core RNA encapsidation signal in murine leukemia virus virions. J. Virol. **71:** 8061–8065.
46. **Muriaux, D., P.-M. Girard, B. Bonnet-Mathonière, and J. Paoletti.** 1995. Dimerization of HIV-1LAI at low ionic strength. An autocomplementary sequence in the 5′ leader region is evidenced by an antisense oligonucleotide. J. Biol. Chem. **270:**8209–8216.
47. **Murti, K. G., M. Bondurant, and A. Tereba.** 1981. Secondary structural features in the 70S RNAs of Moloney murine leukemia and Rous sarcoma viruses as observed by electron microscopy. J. Virol. **37:**411–419.
48. **Osterburg, G., and R. Sommer.** 1981. Computer support of DNA sequence analysis. Comput. Programs Biomed. **13:**101–109.
49. **Paillart, J.-C., E. Skripkin, B. Ehresmann, C. Ehresmann, and R. Marquet.** 1996. A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. Proc. Natl. Acad. Sci. USA **93:**5572–5577.
50. **Paillart, J.-C., L. Berthoux, M. Ottmann, J.-L. Darlix, R. Marquet, B. Ehresmann, and C. Ehresmann.** 1996. A dual role of the putative RNA dimerization initiation site of HIV-1 in genomic RNA packaging and proviral synthesis. J. Virol. **70:**8348–8354.
51. **Paillart, J.-C., E. Westhof, C. Ehresmann, B. Ehresmann, and R. Marquet.** 1997. Noncanonical interactions in a kissing loop complex: the dimerisation initiation site of HIV-1 genomic RNA. J. Mol. Biol. **270:**36–49.
52. **Peterson, R. D., and J. Feigon.** 1996. Structural change in *rev* responsive element RNA of HIV-1 on binding *rev* peptide. J. Mol. Biol. **264:**863–877.
53. **Potts, B. J.** 1990. Mini RT assay, p. 103–106. *In* A. Aldovini and B. D. Walker (ed.), Techniques in HIV research. Stockton Press, New York, N.Y.
54. **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.
55. **Romaniuk, P. J., P. Lowary, H. N. Wu, G. Stormo, and O. C. Uhlenbeck.** 1987. RNA binding site of R17 coat protein. Biochemistry **26:**1563–1568.
56. **Rosen, C. A., A. Terwilliger, J. Dayton, J. G. Sodroski, and W. A. Haseltine.** 1988. Intragenic *cis*-acting ART responsive sequences of HIV-1. Proc. Natl. Acad. Sci. USA **85:**2071–2085.
57. **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 HIV-1 nucleocapsid protein. Proc. Natl. Acad. Sci. USA **90:**5219–5223.
58. **Schlesinger, S., S. Makino, and M. L. Linial.** 1994. *cis*-acting genomic elements and *trans*-acting proteins involved in the assembly of RNA viruses. Semin. Virol. **5:**39–49.
59. **Selden, R. F.** 1987. Transfection using DEAE-dextran, p. 9.2.1.–9.2.6. *In* R. Asubel, R. E. Brent, D. D. Kingston, J. G. Moore, J. A. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology, vol. 1. Greene Publishing Associates and Wiley Interscience, New York, N.Y.
60. **Skripkin, E., J.-C. Paillart, R. Marquet, B. Ehresmann, and C. Ehresmann.** 1994. Identification of the primary site of HIV-1 RNA dimerization *in vitro*. Proc. Natl. Acad. Sci. USA **91:**4945–4949.
61. **Sundquist, W. I., and S. Heaphy.** 1993. Evidence for interstrand quadruplex formation in the dimerization of HIV-1 genomic RNA. Proc. Natl. Acad. Sci. USA **90:**3393–3397.
62. **Valegard, K., J. B. Murray, P. G. Stockley, N. J. Stonehouse, and L. Lijas.** 1994. Crystal structure of an RNA bacteriophage coat protein-operator complex. Nature (London) **371:**623–626.
63. **White, S. A., H. Li, and M. E. Rauch.** 1995. A thermodynamic and mutational analysis of an RNA purine loop as a protein binding site. J. Biomol. Struct. Dyn. **13:**285–299.
64. **Williamson, J. R., M. K. Raghuraman, and T. R. Cech.** 1989. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell **59:** 871–880.
65. **Yang, S., and H. M. Temin.** 1994. A double hairpin structure is necessary for the encapsidation of spleen necrosis virus retroviral RNA. EMBO J. **13:**713–726.
66. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.