Volume 2 · Number 3 · 1992
GENE EXPRESSION

# Sequences within the R region of the long terminal repeat activate basal transcription from the HIV-1 promoter

**Kathleen A. Boris-Lawrie,[1] John N. Brady,[2] and Ajit Kumar[1]**

[1] Department of Biochemistry and Molecular Biology, The George Washington University Medical Center, Washington, D.C. and [2] Laboratory of Molecular Virology, National Cancer Institute, Bethesda, Maryland

**The importance of the R region in basal human immunodeficiency virus type 1 (HIV-1) transcription was addressed by comparing a panel of HIV-1 R region mutants using in vitro and in vivo assays. Using deletion, base substitution mutants, and compensatory mutants, the precise R region sequences essential for basal HIV-1 promoter activity in vitro were mapped to sequences between +17 to +21. Within this regulatory domain, nucleotides +19 and +21 appear to be critical. The effect of these mutations on steady state RNA levels in transfected cells has been analyzed by S1 nuclease protection assay using uniformly labeled probes. Two main conclusions may be drawn from these studies. First, HIV-1 basal transcription is abundant, with the majority of correctly initiated transcripts truncated between sequences +57 to +70. Second, analysis of the compensatory mutants indicates the secondary structure of the nascent R region RNA is not an obligate requirement for the production of the truncated transcripts. Mutations in R region primary sequence that selectively abolish the production of the truncated transcripts in vivo also exhibit reduced promoter activity in vitro. The appearance of high levels of truncated transcripts raise the interesting possibility that—similar to c-*myc*, c-*myb*, and c-*fos*—basal HIV-1 expression is regulated by transcription elongation.**

HIV-1 is the causative agent of AIDS (Barre-Sinoussi et al., 1983; Gallo et al., 1984; Levy et al., 1984). The primary target cell for HIV-1 infection is the CD4+ T lymphocyte (Fauci, 1988). In AIDS patients, the majority of infected T cells are transcriptionally latent (Harper et al., 1986; Zack et al., 1988; Ho et al., 1989; Schnittman et al., 1989), and mitogenic factors that activate T cell transcription stimulate basal transcription, which triggers viral replication (McDougal et al., 1985; Folks et al., 1986; Tong-Starksen et al., 1987; Duh et al., 1989; Harrich et al., 1990). Upon activation of basal transcription, viral proteins Tat and Rev are expressed. Tat protein transactivates the synthesis of all viral transcripts. Rev regulates expression of unspliced and singly spliced viral RNAs, leading to productive viral infection (Cullen and Greene, 1989). The transactivation response element of Tat, TAR, is located within the R region of the LTR (Rosen et al., 1985) between residues +19 to +44, relative to the transcription start site +1 (Hauber and Cullen, 1988; Jakobovits et al., 1988; Feng and Holland, 1988; Selby et al., 1989; Garcia et al., 1989; Berkhout et al., 1989a). The nascent R region RNA forms a stable stem loop structure (Muesing et al., 1987) that encompasses TAR and is essential for Tat function (Berkhout

Received December 24, 1991; revision accepted March 11, 1992.

Correspondence: Ajit Kumar, Department of Biochemistry and Molecular Biology, The George Washington University Medical Center, Washington, DC 20037 Tel (202) 994-3539 Fax (202) 994-8974

 1052-2166/92/0203/215-16$2.00

et al., 1989b; Southgate et al., 1990; Selby and Peterlin, 1990).

The 5′ promoter recognition elements and cognate cellular transcription factors governing HIV-1 basal transcription include the TATA box, which positions the transcription start site and promotes basal transcription (Garcia et al., 1987; Jones et al., 1988; Jakobovits et al., 1988); three GC rich repeats (−45 to −77) that provide binding sites for SP1 (Jones et al., 1986; Leonard et al., 1989; Harrich et al., 1990); and an enhancer element (−78 to −105) that binds several cellular proteins, including EBP-1, NFkB, and the related NFkB family members HIVen86 and p49 (Nabel and Baltimore, 1987; Wu et al., 1988; Shaw et al., 1988; Franza et al., 1988; Bohnlein et al., 1988; Clark et al., 1990; Lee et al., 1990; Schmid et al., 1991). Sequences between −453 and −104 may also contribute to HIV transcription regulation (Zeichner et al., 1991).

The HIV-1 R region has been subjected to rigorous substitution mutagenesis. The primary focus of these studies, however, was the role of R sequences in transactivation by Tat and other transactivators, which required low level basal gene expression in order to maximize responsiveness (Wright et al., 1986; Tong-Starksen et al., 1987; Garcia et al., 1987; Jakobovits et al., 1988; Selby et al., 1989; Garcia et al., 1989; Kliewer et al., 1989; Bohan et al., 1989; Harrich et al., 1990). As a result, the role of the R sequence in basal transcription was not analyzed in detail. Interaction of cellular transcription factors LBP-1 and UBP-1 with sequences downstream of the transcription start site was important for HIV-1 transcription in vivo and in vitro (Jones et al., 1988; Wu et al., 1988). Analysis of substitution mutations within the LBP-1 recognition repeats 5′(T/A)CTGG3′ or 5′CCAG (A/G)3′, using in vitro transcription assays, suggested that LBP-1 activates transcription (Jones et al., 1988). In contrast, purified LBP-1 added to a reconstituted in vitro transcription system conditionally repressed HIV-1 transcription (Kato et al., 1991). Further, Malim et al. (1989) used gel shift analysis to locate three contact sites for cellular protein(s) within sequences −2 to +21. These sequences overlap three LBP-1 recognition sites. In contrast to the results of Jones et al. (1988) and Kato et al. (1991), CAT reporter gene analysis indicated no transcriptional role for these LBP-1 repeats. The different conclusions drawn from these studies concerning the role of LBP-1 in basal transcription suggest that R sequences may mediate multiple regulatory effects.

The synthesis of truncated R region transcripts during basal transcription from the HIV-1 promoter has not been consistently observed (Kao et al., 1987; Hauber and Cullen, 1988; Jones et al., 1988; Toohey and Jones, 1989; Selby et al., 1989; Laspia et al., 1989; Roy et al., 1990; Ratnasabapathy et al., 1990). Truncated transcripts were not observed in conventional in vitro systems (Jones et al., 1988), but were detectable in fractionated HeLa extract or in transcription reactions treated with the anionic detergent, sarkosyl (Toohey and Jones, 1989). Activation or stable production of the truncated transcripts has been postulated to be dependent on HIV-1 promoter sequences and/or the structure of the HIV-1 leader RNA (Toohey and Jones, 1989; Selby et al., 1989; Roy et al., 1990).

In this study, the role of the R sequences in basal transcription from the HIV-1 promoter was evaluated using a panel of R region deletion, base substitution, and compensatory mutants. Analysis of leader deletion mutants by in vitro transcription assays indicated basal promoter activity was abolished in response to deletion of R sequences +3 to +57. Site specific mutations that disrupted R region RNA secondary structure did not alter steady state RNA levels. In contrast, mutation of R sequences +17 to +21 abolished promoter activity. We conclude that the HIV-1 R region contains a promoter element that activates basal transcription.

## Materials and methods

### Expression vector and leader mutations

The expression vector pBennCAT/LM1 was constructed from the HIV-1 LAV strain U3-R region and cloned into pBR322 as an EcoR I-Hind III fragment (Gendleman et al., 1986). A unique Kpn I site was engineered at −2 relative to the transcription start site by replacement of the Pvu II (−17) to Hind III (+82) fragment with a Pvu II (−17) to Hind III (+57) oligonucleotide cassette containing base pair changes at −7 and −2 ($T_{-7}$ to $G_{-7}$ and $T_{-2}$ to $C_{-2}$, LM2). To generate deletion, substitution, and compensatory mutations within the R region, synthetic oligonucleotide cassettes were substituted in the unique Kpn I-Hind III sites (−2 and +57)

of LM2. As illustrated in Figure 1A, sequences +3 to +82, +6 to +82, and +20 to +82 were deleted in LM3, LM4, and LM5, respectively, by substitution of Kpn I/Hind III oligonucleotide cassettes containing residues −2 to +2, −2 to +5, or −2 to +19. To delete residues +3 to +18 (LM7) or +3 to +33 (LM8), LM2 was digested with Kpn I and Bgl II (−2 and +19) or Kpn I and Sac I (−2 and +34) and ligated with the complementary linkers. Substitution and compensatory mutations are illustrated in Figure 1B. Mutations that altered the primary sequence but maintained the secondary structure of the R region RNA were generated by substitution of oligonucleotide cassettes in which pairs of complementary nucleotides that form the predicted RNA secondary structure were switched to the opposite strand (Fig. 1B, TM22, TM24, TM25, TM28, TM32). Mutations that interrupted the RNA secondary structure included TM37 (+18), LM11 (+39, +40, +41), and LM245′ (+17 to +21; Fig. 1B). In LM15, the R sequence was extended to +69 by introducing a Sac I/Hind III (+32 to +69) oligonucleotide cassette into LM2 (Fig. 1A). The complete R region of the LTR (98 nt) and U5 sequences to +126 have been maintained in the CD12 plasmid (Siekevitz et al., 1987). All constructs were confirmed by DNA sequencing using the dideoxy chain termination method with a DNA primer complementary to CAT sequences between +49 to +80.

### In vitro transcription assays

Whole-cell extracts for in vitro transcription assays were prepared as described by Manley et al. (1980). Transcription reactions (15 μl) contained 7.5 μl of whole cell HeLa extract, 50 ng supercoiled template DNA resuspended in $H_2O$, 0.5 μg/ml poly(dI:dC), and 1 mM nucleoside triphosphates (Pharmacia). Reactions were incubated at 30°C for 30 minutes and then treated with 1 unit of RQ1 DNase (Promega Biotech) added to the transcription reaction in 15 μl DNase buffer (10 mM Tris pH 7.5, 500 mM NaCl, 50 mM $MgCl_2$, 2 mM $CaCl_2$). After incubation at 30°C for 15 minutes, the reaction was stopped by the addition of 255 μl stop buffer (0.2% SDS, 10 mM Tris, pH 7.5, 100 mM NaCl), and 30 μl 3M NaOAc, pH 5.2. The reaction was extracted with an equal volume acid phenol: chloroform:isoamyl alcohol (25:24:1) and chloroform:isoamyl alcohol and precipitated with 2.5 volumes of 100% ethanol. The in vitro transcripts generated from supercoiled template DNA were analyzed by the S1 nuclease protection assay described below and separated on 6% denaturing polyacrylamide gels with $^{32}P$ end-labeled Msp I-digested pBR322 marker DNA. Run-off transcription reactions were identical, except that the template DNA was truncated within the CAT gene at the EcoR I site and transcribed in the presence of [α-$^{32}P$] UTP (400 Ci/mmol; Amersham) without cold UTP. The reactions were extracted, precipitated by ethanol, and analyzed on 6% denaturing polyacrylamide gels. The results of the in vitro transcription assays were consistent in at least five repetitions of each assay in four separately prepared whole-cell extracts. In the control experiments comparing RNA isolation techniques using [α-$^{32}P$]-labeled R region and R region-CAT transcripts, run-off transcripts were expressed from the T7 promoter upstream of R region-CAT sequences. The T7 template was truncated at either the Hind III or the Nco I site to generate 70 nt or 632 nt run-off transcripts, respectively. The labeled transcripts were isolated from 6% denaturing polyacrylamide gels and eluted in elution buffer (500 mM $NH_4C_2H_3O_2$, 10 mM $MgC_2H_3O_2$, 1 mM EDTA).

### Generation of $^{32}P$-uniformly labeled probes

Single-stranded antisense DNA probes were generated using cycles of primer extension in the presence of [α-$^{32}P$]dCTP. Plasmid template DNA (500 ng) was cleaved at −17 with Pvu II and combined with a 31-base oligonucleotide antisense primer (1 μM) complementary to CAT sequences (CAT nucleotides 49 to 80), 1X standard Perkin-Elmer-Cetus PCR reaction buffer, 20 μM dGTP, dATP, dTTP and 12 μl [α-$^{32}P$] dCTP (400 mCi/mmol; Amersham), and Taq I polymerase at 25 U/ml and topped with one drop of mineral oil. Cycles of plasmid denaturation, primer:template annealing, and primer extension were achieved in a Perkin-Elmer-Cetus DNA Amplifier for 30 cycles of 94°C for 1 minute, 37°C for 1 minute, and 72°C for 2 minutes. The reaction was extracted by chloroform, precipitated by ethanol, and separated on a 6% denaturing polyacrylamide gel. The labeled product was eluted from the gel slice in elution buffer. For each template, the predicted product contained 17 bases of 5′ promoter sequence, template R sequence, and 80 bases of CAT sequence.

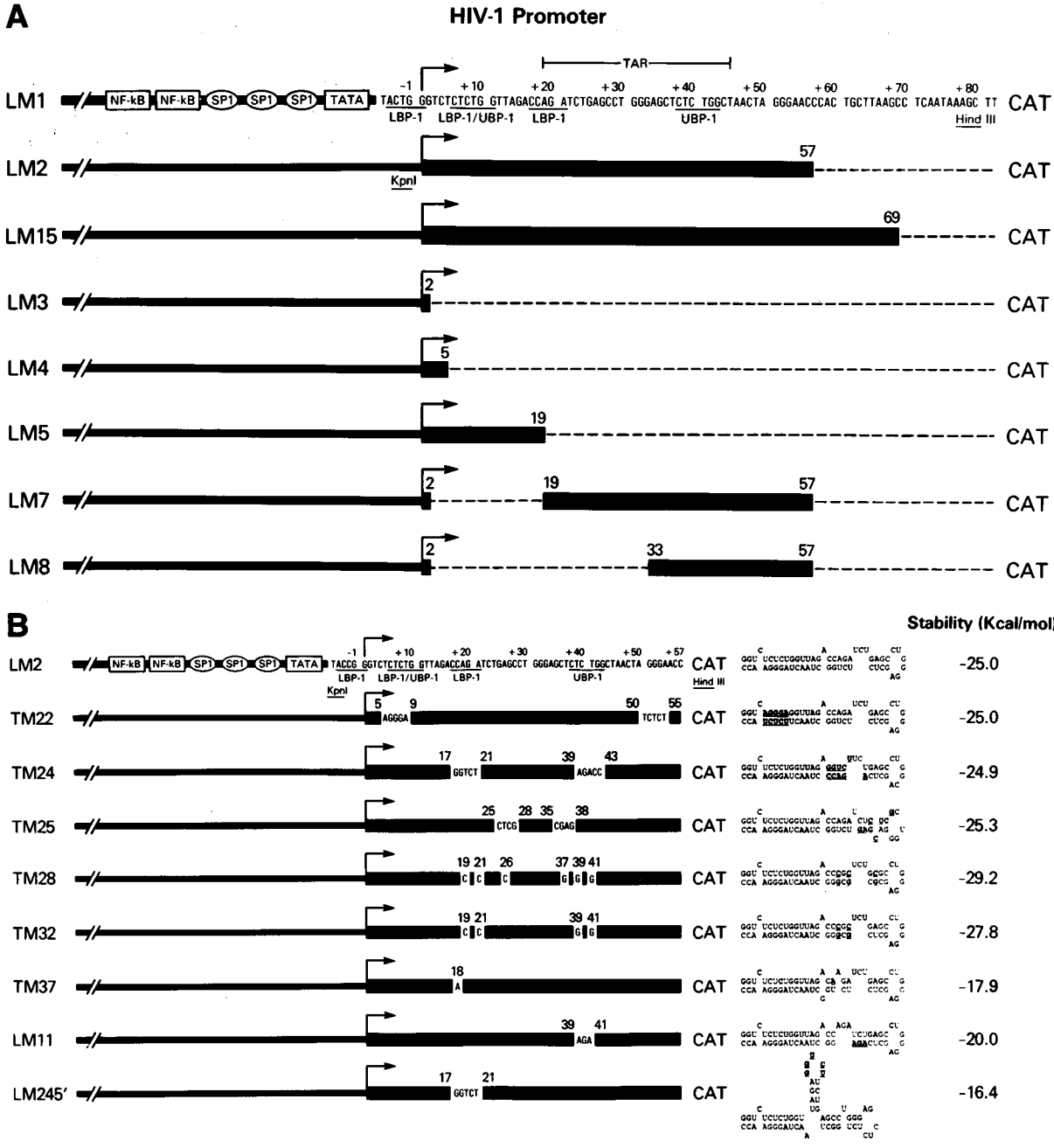


**Figure 1.** The HIV-1 R region mutations. **A.** HIV-1 LTR CAT R region deletion mutations. The HIV-1 LTR-CAT expression plasmid LM1 contains R sequences +1 to +82 inclusive of the TAR region (+19 to +44). The transcription start site (+1) is designated by an arrow. In LM2, R sequences (+1 to +57) are designated by a black bar and are labeled with the endpoint nucleotide. Subsequent deletions were constructed in LM2 by substitution of synthetic mutated oligonucleotide cassettes in the unique Kpn I/Hind III sites. The recognition sites for LBP-1/UBP-1 are underlined. Also indicated are the recognition sites for the restriction endonucleases Hind III and Kpn I, enhancer factor NF-kB (nuclear factor-kB), the SP1 binding sites, TATA box, and the chloramphenicol acetyltransferase (CAT) gene. **B.** HIV-1 R region substitution mutations. Substitution mutations in LM2 were constructed by substitution of synthetic oligonucleotides into the unique Kpn I/Hind III sites. Boundaries of mutated sequences and the altered residues are indicated. The predicted nascent RNA structure and free energy (kcal/mol) for each mutant (Zucker, 1987) are shown in the right panel.

As shown in Figure 2A, using the LM1 template, the 179 nt probe contained 17 bases of 5′ promoter sequence, 82 bases of R sequence, and 80 bases of CAT. LM2 probe was 154 nt, containing 17 bases of 5′ promoter sequence, 57 bases of R sequence, and 80 bases of CAT. The predicted homologous RNA protection products were 17 nt shorter: 162 nt for LM1 and 137 nt for LM2.

### Transient transfection assays

HeLa S cells were cultured in Dulbecco's modified Eagle medium containing 10% fetal calf serum. To prepare for DNA transfection, cells were grown to 50% confluence in 10 cm dishes. Test DNA, carrier DNA, and a control DNA—pON1, used to evaluate transfection efficiency (Spaete and Mocarski, 1985)—were combined with equal volumes of 0.25 M $CaCl_2$ and 2X BBS buffer (50 mM BES [N,N-bis(2-Hydroxyethyl)-2-aminoethanesulfonic acid; Sigma], 280 mM NaCl, 1.5 mM $Na_2HPO_2$) and added dropwise to the cells. After incubation overnight in 3.5% $CO_2$, the DNA precipitate was washed from the cells with serum-free media. Serum-supplemented media was added, and the cells were incubated in 5% $CO_2$. Cells were harvested 48 hours after transfection. The transfection assays were repeated at least five times with separately prepared DNA samples.

### RNA preparation

Transfected HeLa cells (6 μg LTR CAT DNA per $10^6$ cells) were rinsed, harvested in cold PBS, and pelleted at 3000 RPM and 4°C for 1 minute in an Eppendorf centrifuge. To prepare cytoplasmic RNA, cells were resuspended on ice in cell membrane lysis buffer (10 mM Tris pH 8.4, 150 mM NaCl, 1.5 mM $MgCl_2$) and 0.5% Nonidet-P40 (NP40, Sigma). The suspension was mixed, incubated on ice for 20 minutes, and centrifuged for 2 minutes at 3000 RPM and 4°C to pellet nuclei. Supernatant containing cytoplasmic RNA was extracted with phenol:chloroform:isoamyl alcohol and precipitated with 2 volumes of ethanol. Total cellular RNA was prepared by lysis of the cells with 4 M guanidine isothiocyanate and cesium chloride density centrifugation (Chirgwin et al., 1979).

### Analysis of steady-state RNA

S1 nuclease protection analysis was performed using 20 μg HeLa cellular RNA resuspended in 20 μl hybridization buffer (40 mM PIPES [piperazine-N,N′-bis(2-ethanesulfonic acid)] pH 6.4, 1 mM EDTA, 0.2 M NaCl, 80% deionized formamide) containing the probe ($1 \times 10^5$ CPM/reaction). The samples were denatured at 85°C for 10 minutes and allowed to hybridize for 12 hours at 37°C. The samples were digested with 300 units S1 nuclease per reaction in 180 μl digestion buffer (30 mM NaOAc pH 4.6, 0.1 mM NaCl, 20 mM $ZnSO_4$) at room temperature for one hour. Following ethanol precipitation, the protected transcripts were separated on 8% sequencing gels and subjected to autoradiography on preflashed film using standard procedures. Transcript levels were quantitated by excision of the bands and liquid scintillation counting. Transfection efficiency was determined to be equivalent by hybridization of an aliquot of the cellular RNA to [$^{32}$P]-end-labeled pON1 probe (data not shown).

## Results

### Construction of R region mutants

The parental CAT expression vector, pBennCAT/LM1, was constructed from the HIV-1 LAV strain U3-R region and cloned into pBR322 as an EcoR I-Hind III fragment (Gendleman et al., 1986). As summarized in Figure 1A, pBennCAT/LM1 contained R sequences +1 to +82. A unique Kpn I site was generated at −2 by replacement of the Pvu II (−17) to Hind III (+82) fragment with a (−17) Pvu II to (+57) Hind III oligonucleotide cassette containing substitutions at nucleotides $T_{-7}$ and $T_{-2}$ to $G_{-7}$ and $C_{-2}$ (LM2). A panel of deletion, substitution, and compensatory mutations was generated in LM2 by substitution of synthetic Kpn I-Hind III oligonucleotide cassettes, as listed in Figure 1 (A and B). Briefly, R region deletions eliminated residues +3 to + 82 (LM3), or +6 to +82 (LM4), +20 to +57 (LM5), +3 to +18 (LM7), or +3 to +32 (LM8) (Fig. 1A). Substitution mutations that altered R primary sequence, but maintained the predicted nascent RNA secondary structure, extended between nucleotides +5 and +55 (Fig. 1B). In these mutants, pairs of nucleotides were switched from the corresponding complementary strands of RNA, and the predicted free energy values of the RNA secondary structures remained similar to LM2 (Fig. 1B, TM22, TM24, TM25, TM28, TM32). The compensatory sub-

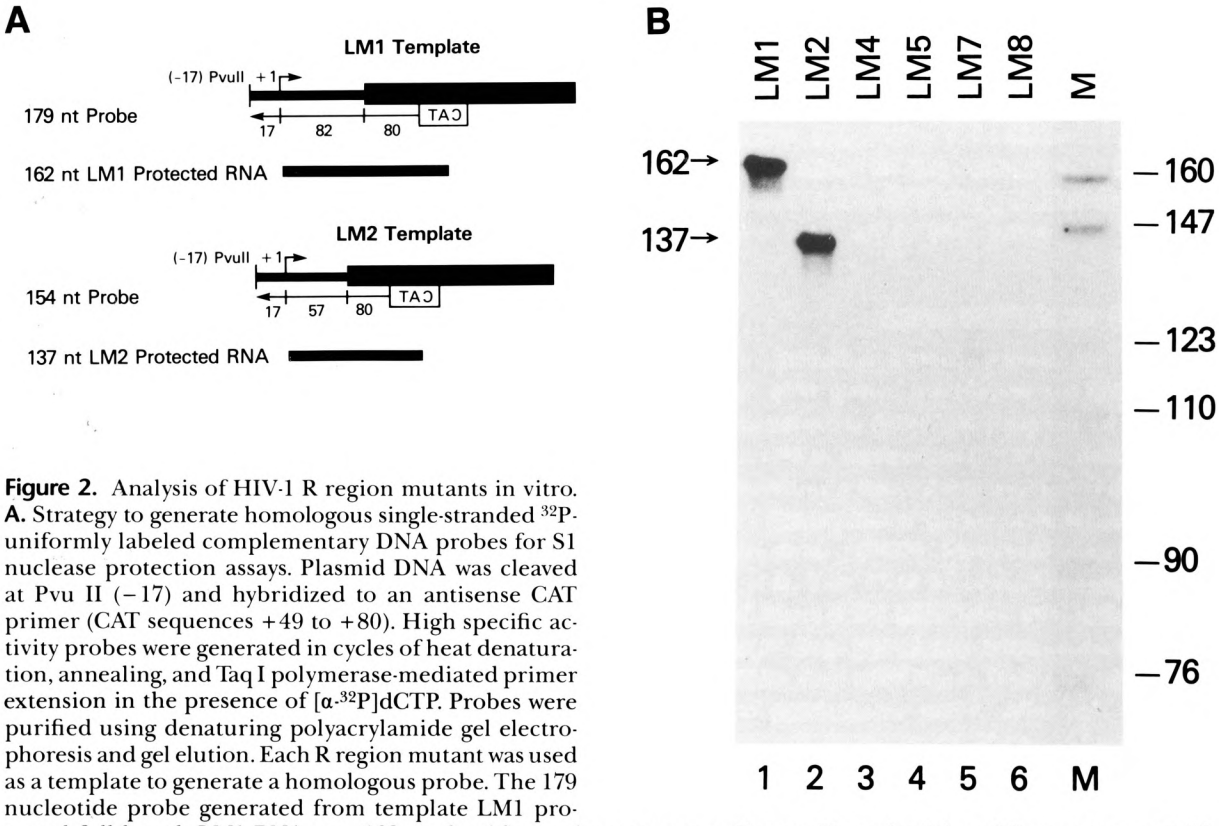


**Figure 2.** Analysis of HIV-1 R region mutants in vitro. **A.** Strategy to generate homologous single-stranded $^{32}$P-uniformly labeled complementary DNA probes for S1 nuclease protection assays. Plasmid DNA was cleaved at Pvu II (−17) and hybridized to an antisense CAT primer (CAT sequences +49 to +80). High specific activity probes were generated in cycles of heat denaturation, annealing, and Taq I polymerase-mediated primer extension in the presence of [α-$^{32}$P]dCTP. Probes were purified using denaturing polyacrylamide gel electrophoresis and gel elution. Each R region mutant was used as a template to generate a homologous probe. The 179 nucleotide probe generated from template LM1 protected full length LM1 RNA as a 162 nucleotide product corresponding to 82 nucleotides of R sequence and 80 nucleotides of CAT sequence. The 154 nucleotide LM2 probe protected 137 nucleotides of full length LM2 RNA corresponding to 57 nucleotides of R sequence and 80 bases of CAT. **B.** R sequences were essential for in vitro transcription from the HIV-1 promoter. DNA templates (50 ng) were transcribed in vitro using HeLa whole cell extracts, treated with DNase, and analyzed by the S1 protection assay using the homologous R region probe. Lanes and predicted transcription size are as follows. Lane 1: LM1, 162 nt; lane 2: LM2, 137 nt; lane 3: LM4, 85 nt; lane 4: LM5, 99 nt; lane 5: LM7, 116 nt; lane 6: LM8, 108 nt; lane M: $^{32}$P-labeled pBR322 Msp I digested DNA markers.

stitutions in TM22 interrupted the LBP-1/UBP-1 motif between +7 and +11. The base changes in TM24 disrupted the putative LBP-1 site between +17 and +21 and repositioned the motif to the opposite strand. The four-base switch in TM25 included a one-base change in the LBP-1 motif at +28 to +32 and three flanking nucleotides. Substitution mutations that were not compensatory for RNA structure included TM37, LM11, and LM245′ (Fig. 1B). In LM15, the R region was extended to +69 by inserting into LM2 a Sac I/Hind III oligonucleotide cassette containing R sequences from +32 to +69 (Fig. 1A). The complete R region (98 nt) and U5 sequences to +126 of HIV-1 IIIB were maintained in plasmid CD12 (not shown; Siekevitz et al., 1987). All constructs were confirmed by DNA sequencing using the dideoxy chain termination method with a DNA primer complementary to CAT sequences between +49 to +80.

### R sequences are essential for HIV-1 promoter activity in vitro

To evaluate downstream sequences important for HIV-1 promoter activity, the transcriptional activity of HIV-1 R deletion mutants was assayed in HeLa whole-cell extracts. Each template DNA was transcribed in vitro (Manley et al., 1980) and analyzed by S1 nuclease protection following hybridization to a uniformly-labeled, single-stranded DNA probe (Fig. 2A). Template DNA containing R residues +1 to +82 (LM1) or +1 to +57 (LM2) exhibited high level transcription of the predicted full-length transcript (Fig. 2B, lanes 1 and 2), indicating the two base substitutions and deletion of R sequences +58 to +82

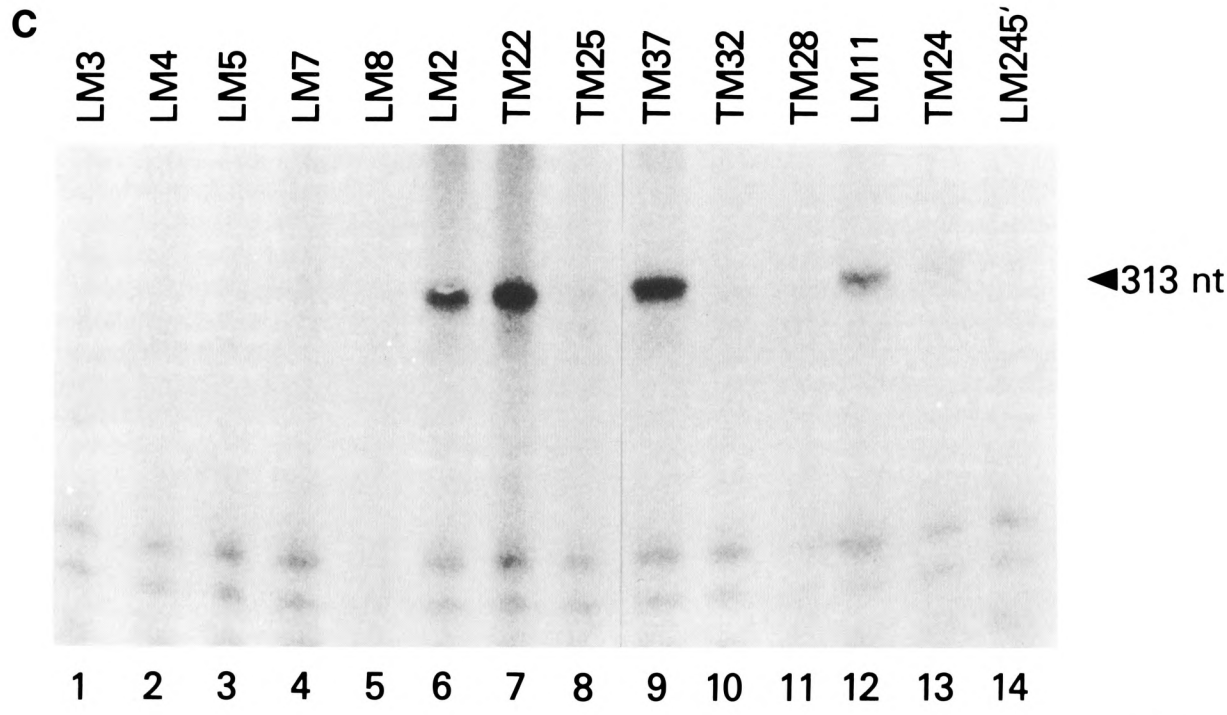


**Figure 2 (continued). C.** HIV-1 promoter activity in vitro was abolished in R region substitution mutants. Template DNA was truncated at the EcoR I site, and run-off transcripts (313 nt) were generated in vitro using HeLa whole cell extracts. Lane 1: LM3, 258 nt; lane 2: LM4, 261 nt; lane 3: LM5, 275 nt; lane 4: LM7, 292 nt; lane 5: LM8, 280 nt; lane 6: LM2, 313 nt; lane 7: TM22, 313 nt; lane 8: TM25, 313 nt; lane 9: TM37, 313 nt; lane 10: TM32, 313 nt; lane 11: TM28, 313 nt; lane 12: LM11, 313 nt; lane 13: TM24, 313 nt; lane 14: TM245′, 313 nt.

in LM2 have no effect on transcriptional activity. In contrast, deletion of R sequences +6 to +57 (LM4), +20 to +57 (LM5), +3 to +18 (LM7), and +3 to +32 (LM8) abolished promoter activity in vitro (Fig. 2B, lanes 3–6). These results confirm that R sequences between +3 to +57 are critical for the transcriptional activity of the HIV-1 promoter in vitro.

### Substitution mutations abolish HIV-1 promoter activity in vitro

To define more precisely the R sequences essential for basal transcription in vitro, template DNAs containing substitution mutations within the R region were cleaved with EcoR I and analyzed in run-off in vitro transcription assays. The in vitro assays were repeated at least five times in each of four independently prepared whole-cell extracts. All HIV-1 transcripts were sensitive to 1 μg/ml α-amanitin. As shown in Figure 2C (lane 6), the control template LM2 was active in the run-off transcription assay expressing high levels of the predicted 313 nucleotide transcript. The deletion mutants (Fig. 2C, lanes 1–5; for predicted transcript size, see legend) were inactive, indicating the loss of promoter activity. The level of the 313 nucleotide run-off transcript from the noncompensatory base substitution mutants TM37 and LM11 was similar to that of the wild-type LM2 (Fig. 2C, lanes 6, 9, 12). The compensatory mutant TM22 exhibited wild-type transcriptional activity (Fig. 2C, lane 7). Another compensatory mutant, TM25, produced inconsistent results and was inactive in the assay shown here (Fig. 2C, lane 8). The absence of promoter activity from LM245′ (Fig. 2C, lane 14) and TM24 (Fig. 2C, lane 13) suggested that sequences +17 to +21 were required for promoter activity. Consistent with these nucleotides being required for promoter activity in vitro, transcriptional activity was abolished from templates TM32 (Fig. 2C, lane 10) and TM28 (Fig. 2C, lane 11), while LM11 (Fig. 2C, lane 12) exhibited wild-type level transcription. Thus, the R region sequence between +17 to +21

constitutes an important basal promoter element, and nucleotides +19 and +21 appear to be critical for optimum HIV-1 basal transcription.

### R region deletions significantly reduce steady-state HIV-1 RNA levels

Since HIV-1 R region sequences were requisite for basal transcription in vitro, the role of these sequences in the abundance of steady-state RNA was assayed by S1 nuclease protection analysis. Analysis of steady-state RNA from cells transfected with LM1 and LM2 indicated strong basal transcription from the HIV-1 promoter, as measured by the abundance of incompletely elongated HIV-1 transcripts (Fig. 3). In cells transfected with LM1, less than 1% of the transcripts initiating at +1 were full-length R region-CAT RNA. Transcripts initiated at +1 and truncated 57 to 70 nt downstream represented a 232 molar excess compared to the abundance of full-length transcripts (Fig. 3, lane 6). In cells transfected with LM2 (Fig. 3, lane 1), the levels of full-length transcript and truncated transcript were consistently similar to those of the wild-type LM1 and suggested that neither the two-base substitutions in LM2 nor the deletion of sequences downstream of +57 affected steady-state RNA level.

In contrast to the strong basal transcription detected from LM1 and LM2, when cells were

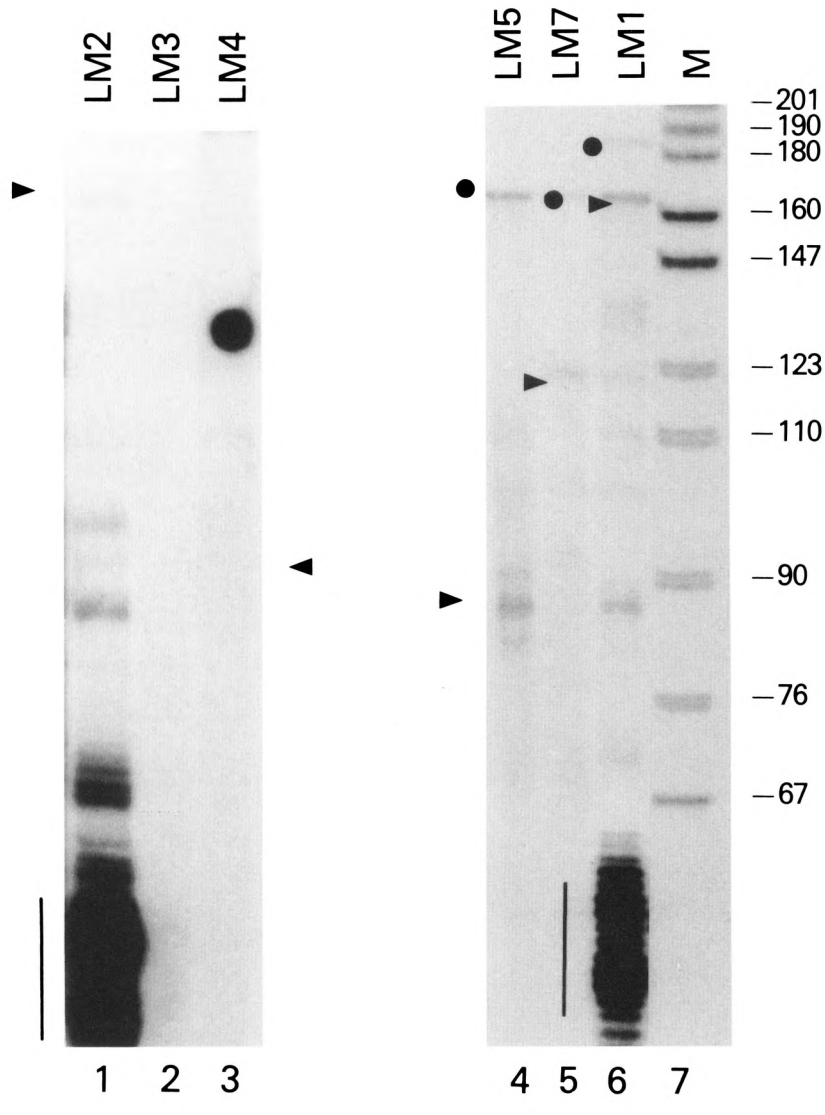


**Figure 3.** Steady-state HIV-1 RNA levels were significantly reduced in R region deletion mutants. S1 nuclease protection analysis of RNA from cells transfected with the R region deletion mutants. The truncated transcripts are indicated by bars, full-length transcripts by arrows, and undigested probe by dots. **A.** RNA from transfected cells (20 μg) was hybridized to the homologous probes. Lane 1: LM2; lane 2: LM3; lane 3: LM4. **B.** LM2 probe was hybridized to each RNA sample, except for LM1, which was hybridized to LM1 probe. Lane 4: LM5; lane 5: LM7; lane 6: LM1; lane 7: LM1; lane M: $^{32}$P-labeled pBR322 Msp I DNA markers.

transfected with R mutants containing deletion of residues +3 to +57 (LM3) or +6 to +57 (LM4), no full-length or truncated HIV-1 transcripts were detected (Fig. 3, lanes 2 and 3). In cells transfected with LM5 (+1 to +19) or LM7 (+19 to +57), full-length transcripts were observed in low abundance (Fig. 3, lanes 4 and 5). No truncated transcripts 57 to 70 nt in length or shorter transcripts (i.e., 19 nt for LM5 or 38 nt for LM7) were detected in several repetitions of the S1 analysis using the homologous probes (data not shown). The absence of truncated transcripts encoded by the R deletion mutants correlated with the lack of in vitro transcriptional activity from these mutants.

### Wild-type RNA secondary structure is not an obligate requirement for the production of the truncated transcripts

The absence of HIV-1 truncated transcripts could result from diminished nascent RNA stability. To analyze the contribution of secondary structure and primary sequence to in vivo basal transcription, R region substitution mutations that were compensatory for RNA secondary structure were utilized. The panel of compensatory substitution mutant DNAs were transfected into HeLa cells, and steady-state RNA was analyzed by S1 nuclease protection assays using homologous probes. S1 nuclease protection analysis of steady state RNA from cells transfected with TM22, TM24, and TM25 detected neither full-length (137 nt) nor truncated transcripts (57–70 nt) (Fig. 4, lanes 1, 2, 3). The ability of these probes to hybridize and protect the cognate RNA was confirmed in a control experiment by hybridization to the homologous RNAs (data not shown). When these constructs were transfected into HeLa cells constitutively producing Tat, no RNA was detected in response to Tat (data not shown). Since the predicted secondary structure of TM22, TM24, and TM25 nascent RNA was similar to the wild-type LM2 control RNA, these results suggested that primary sequence, independent of RNA secondary structure, regulated steady-state RNA levels.

In contrast to the compensatory base substitution mutations in TM22, TM24, and TM25, the noncompensatory mutation in TM37 (C to A transversion at +16) altered the structure of the nascent leader RNA and reduced the predicted free energy from −25 kcal/mol (LM2) to −17.9 kcal/mol (Fig. 1B). However, wild-type levels of full-length and truncated transcripts were detected from cells transfected with TM37 (Fig. 4, lane 8). The noncompensatory mutant LM11 contained three base substitution mutations (+39, +40, +41) that reduced predicted RNA free energy to −20 kcal/mol (Fig. 1B). Similarly, in the analysis of LM11 steady state RNA, wild-type levels of full-length and truncated transcripts were observed (Fig. 4, lane 5). These results suggested that maintenance of the wild-type RNA secondary structure was not obligatory for the production of the truncated transcripts.

We next extended our analysis to compensatory primary sequence mutations. The compensatory mutant TM28 contained base substitutions in three sets of paired nucleotides (+19 and +41, +21 and +39, +26 and +37). Analysis of steady-state RNA revealed full-length transcripts (137 nt) but no truncated transcripts (Fig. 4, lane 4). Similarly, in response to compensatory mutations at +19 and +41, +21 and +39 (TM32), no truncated transcripts were observed in steady-state RNA, and a low level of the full-length transcript was exhibited (Fig. 4, lane 7). These results suggest that the primary sequence of the R region was important for the production of the truncated transcripts.

The selective loss of the truncated transcripts in TM28 and TM32 could be due to mutations in nucleotides +19 and +21 or nucleotides +39 and +41. As described above, LM11 contained mutations in nucleotides +39, +40, +41 that were not compensatory for RNA structure (Fig. 1B). LM11 produced wild-type levels of full-length and truncated transcripts (Fig. 4, lane 5). This result suggested that the mutated 3′ segment of TM32 (+39 and +41) did not influence steady-state RNA level, and that the selective loss of the truncated transcripts in TM28 and TM32 was due to the mutation in R sequences +19 and +21. Analysis of steady-state RNA from LM245′ (+17 to +21) revealed no full-length or truncated transcripts (data not shown), which agreed with the in vitro data suggesting that R sequences between +17 to +21 are essential for high level basal promoter activity.

### Mapping of HIV-1 truncated transcript termini

These results have shown that HIV-1 R region sequences were essential for basal transcription in vitro and correlated with the production of

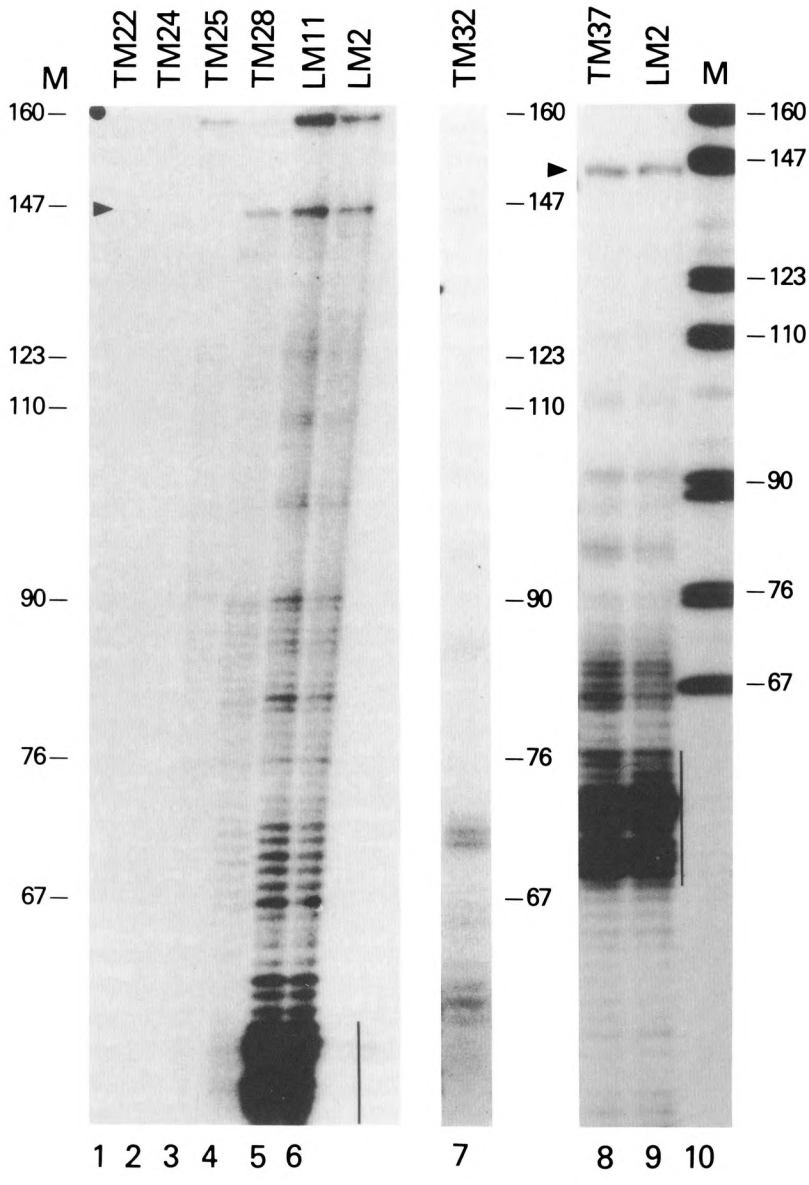


**Figure 4.** R region substitution mutations selectively abolished the truncated transcripts. S1 nuclease protection analysis of steady-state RNA from R region substitution mutants. Cellular RNA (20 μg) and the homologous probe were hybridized, treated with S1 nuclease, and analyzed on 8% sequencing gels. Truncated transcripts are labeled with bars, full-length transcripts with arrows, and undigested probe with a dot. Lane 1: TM22; lane 2: TM24; lane 3: TM25; lane 4: TM28; lane 5: LM11; lane 6: LM2; lane 7: TM32; lane 8: TM37; lane 9: LM2; lane 10: $^{32}$P-labeled pBR322 Msp I DNA markers.

truncated transcripts in steady-state HIV-1 RNA. The truncated transcripts represented the majority of transcription products from the HIV-1 basal promoter. Compared to the full-length transcripts, the truncated transcripts represent a molar excess of 232:1. To map the transcription start site (+1) of the truncated transcripts, the transcript sizes protected by two uniformly-labeled probes (+1 and +19) were compared. The first probe extended from −17 to +137 (+1 probe) and protected LM2 full-length RNAs as 137 nt transcripts and the truncated RNAs as 57 to 70 nt fragments (Fig. 5A, lane 1). In comparison, an LM2 probe complementary to residues +19 to +137 protected a 118 nt fragment corresponding to full-length transcripts (Fig. 5A, lane 2). Likewise, the truncated transcripts were 18 nt shorter (39 to 52 nt; Fig. 5A, lane 2). A sequencing ladder run in parallel was used to determine precisely transcript size (not shown). From these results we concluded that the truncated and full-length transcripts initiated from the same RNA start site (+1) and were terminated between 57 to 70 nucleotides downstream of +1.

To address the effect of HIV-1 leader residues

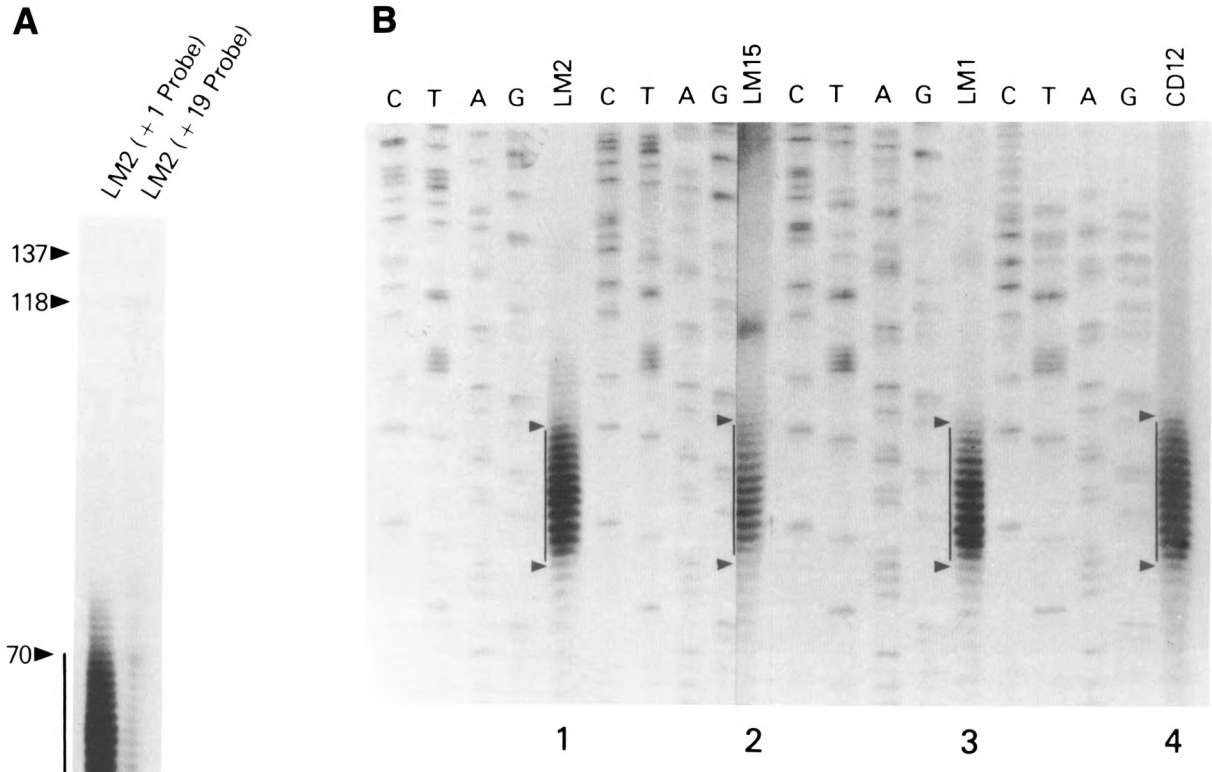


**Figure 5.** Mapping of the termini of the truncated transcripts. **A.** Correctly initiated HIV-1 transcripts terminated 57 to 70 nt downstream of the transcription start site and initiated at +1. The termini of the truncated transcripts were mapped in LM2 RNA using two uniformly-labeled probes (+1 probe and +19 probe). LM2 +1 probe extended from −17 to +137, as illustrated in Figure 2A, and protected the transcripts from +1 to +137 (lane 1, arrow) and truncated transcripts 57 to 70 nt (lane 1, bar). To generate the +19 probe, LM2 template DNA was truncated at Bgl II (+19). The +19 probe protected the 5′ end of R region-CAT transcript and truncated transcripts beginning 19 nt downstream of +1 (lane 2, bar). **B.** The positioning of the truncated transcripts was independent of R sequences downstream of +57. S1 nuclease protection analysis of steady-state RNA from cells transfected with R region constructs containing extended 3′ termini. Each RNA was hybridized to the homologous probe and was compared to the homologous sequencing ladder, as shown to the left of each S1 analysis. Lane 1: LM2 (R sequences +1 to +57); lane 2: LM15 (+1 to +69); lane 3: LM1 (+1 to +80); lane 4: CD12 (+1 to +126).

downstream of +57 on the positioning of the truncated transcripts, the sizes of the truncated transcripts expressed in constructs containing R sequences +1 to +57 (LM2), +1 to +69 (LM15), +1 to +82 (LM1), and +1 to +126 (CD12) were compared (Fig. 5B). HeLa cells were transfected, and cellular RNA was prepared by the NP40 method, hybridized to the homologous R region probe, and treated with S1 nuclease. No significant differences in the level of basal transcripts were detected in response to extension of R sequence downstream of +57 (Fig. 5B, lanes 1–4). Most striking, however, was that the positioning of the truncated transcripts was identical among the constructs. By comparison to the homologous sequencing ladders adjacent to the protected RNAs, the 3′ ends of the LM2 truncated RNAs (Fig. 5B, lane 1) corresponded to sequence 57 to 70 nucleotides downstream of the transcription start site (+1). Since the R region in LM2 extended to +57, the 3′ termini of LM2 truncated transcripts corresponded

to nucleotides 1 to 13 of CAT gene sequence. Similar to LM2, the truncated transcripts expressed from LM15 (+1 to +69; Fig. 5B, lane 2) mapped to nucleotides +57 to +70, the 12 terminal nucleotides of the R sequence and the first nucleotide of CAT sequence. The truncated transcripts expressed from LM1 (+1 to +82; Fig. 5B, lane 3) also terminated at nucleotides +57 to +70, mapping entirely within the HIV-1 R sequence. Similarly, truncated transcripts expressed from CD12 (+1 to +126; Fig. 5B, lane 4) terminated within the R region at nucleotides +57 to +70 downstream of the transcription start site. These data indicated that R sequences +58 to +126 did not influence the level of transcription from the HIV-1 LTR or the positioning of the truncated RNAs. Thus the positioning of the truncated R region transcripts was determined by signals upstream of +58.

### Efficient isolation of the HIV-1 truncated transcripts is dependent on the method of RNA isolation

Since the truncated transcripts were not consistently reported in previous studies, we analyzed the efficiency of recovery of the HIV-1 truncated transcripts by two extraction procedures: guanidine isothiocyanate/cesium chloride gradient separation of total cellular RNA (GSCN/CsCl; Chirgwin, 1979) and NP40 extraction of RNA (see Materials and Methods). In the RNA samples prepared from cells transfected with LM1 or LM2, abundant 57 to 70 nt truncated transcripts were recovered using the NP40 procedure (Fig. 6A, lanes 1 and 3). In contrast, the truncated transcripts were absent in RNA samples isolated by the GSCN/CsCl method (Fig. 6A, lanes 2 and 4). Full-length transcripts were recovered by both methods (Fig. 6A, lanes 1–4). These results suggested either that the truncated RNAs were not recovered in the GSCN/CsCl procedure or that the 57–70 nt RNAs represent breakdown products of the full-length transcripts incurred during the NP40 RNA extraction procedure. To investigate these possibilities, HeLa cells were supplemented with in vitro transcribed $^{32}$P-labeled R region transcript (70 nt in length), and cellular RNA was extracted by the NP40 or GSCN/CsCl gradient procedures. The 70 nt R transcript was not recovered using the GSCN/CsCl procedure (Fig. 6B, lanes 4 and 5), but was recovered using the NP40 procedure (Fig. 6B, lanes 2 and 3). To examine the possibility that the truncated transcripts arose from breakdown of the full-length R region-CAT transcript, we added $^{32}$P-labeled R region-CAT transcript (632 nt in length) to HeLa cells and isolated RNA by the two extraction methods. The 632 nt R region-CAT RNA was fully intact after purification by both the NP40 and GSCN/CsCl isolation procedures (Fig. 6C, lanes 1–4). We concluded that quantitative recovery of authentic HIV-1 transcription products was provided using the NP40 procedure and that the 57–70 nt HIV-1 RNAs were not produced by degradation of longer RNA species.

## Discussion

Our analysis of a panel of substitution and compensatory mutants suggests that sequences within the R region are required for HIV-1 basal transcription. Compensatory site-directed mutations in TM32 or TM28 abolished synthesis of the truncated transcripts in vivo. TM32 and TM28 were inactive in the in vitro transcription assay as well, suggesting these mutations affected the steady-state level of the truncated transcripts by reducing transcriptional activity of the HIV-1 promoter. The base substitutions in mutant LM11 overlapped those in TM32 and TM28. LM11 exhibited wild-type transcriptional activity in the in vitro assays and suggested that the loss of promoter activity in TM32 and TM28 was due specifically to mutations at +19 and +21. Consistent with this result, LM245′, which contained mutations in nucleotides +17 to +21, was inactive in vitro and in vivo. In our hands, mutations that altered the predicted secondary structure of the nascent RNA, TM37 and LM11, did not decrease transcription. However, previous investigations using compensatory R region mutants have proposed a role for the nascent R region RNA secondary structure in the generation of the truncated transcripts (Selby et al., 1989; Roy et al., 1990a). It is possible, therefore, that both primary sequence and secondary structure contribute to the expression of truncated transcripts. This interpretation is consistent with the recent observation that insertion of HIV-1 sequences −5 to +80 downstream of selected RNA polymerase II and III promoters induced expression of truncated tran-

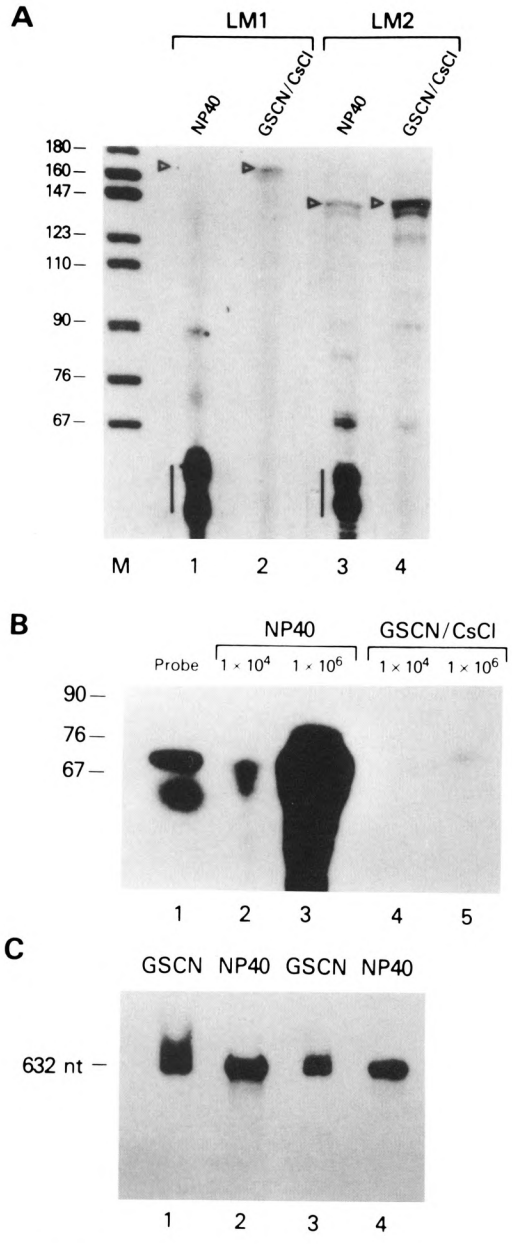


**Figure 6.** Truncated transcripts were present in cellular RNA prepared by NP40 extraction but not by GSCN/CsCl extraction. **A.** S1 nuclease protection analysis of LM1 and LM2 RNA prepared using the NP40 procedure or the GSCN/CsCl gradient method. RNA was isolated from duplicate HeLa cell transfections of LM1 and LM2 DNA, and 20 μg was hybridized with equal molar amounts of LM1 and LM2 uniformly-labeled probes. Protected full-length transcripts are designated by arrows, and the truncated R region transcripts are designated by bars. Lane M: $^{32}$P-labeled markers; lane 1: LM1 RNA isolated by the NP40 method, 162 nt R region-CAT transcript, and 57 nt–70 nt truncated transcripts; lane 2: LM1 RNA isolated by the GSCN/CsCl gradient method, 162 nt transcript; lane 3: LM2 RNA isolated by the NP40 method, 137 nt R region-CAT transcript, and 57–70 nt truncated transcripts; lane 4: LM2 RNA prepared by the GSCN/CsCl gradient method, 137 nt R region-CAT transcript. **B.** Comparison of the recovery of 70 nt $^{32}$P-labeled R region transcript from NP40 and GSCN/CsCl RNA isolation procedures. Lane 1: untreated 70 nt R region transcript; lane 2: R region transcript recovery in NP40 RNA from HeLa cells supplemented with $1 \times 10^4$ CPM R region transcript; lane 3: recovery in NP40 RNA from cells supplemented with $1 \times 10^6$ CPM R region transcript; lane 4: recovery in GSCN/CsCl RNA from cells supplemented with $1 \times 10^4$ CPM R region transcript; lane 5: recovery in GSCN/CsCl RNA from cells supplemented with $1 \times 10^6$ CPM R region transcript. **C.** Comparison of the stability of 632 nt R region-CAT transcript in NP40 or GSCN/CsCl extracted and purified RNA. HeLa cells were supplemented with $^{32}$P-labeled R region-CAT transcript (632 nt), and cellular RNA was extracted. Lanes 1 and 3: R region CAT transcript recovered from GSCN/CsCl extracted RNA; lanes 2 and 4: R region CAT transcript recovered from NP40 extracted RNA.

scripts in addition to the transcripts characteristically expressed from the promoter (Ratnasabapathy et al., 1990).

The results of this study suggest that analysis of HIV-1 transcription requires quantitation of RNA using probes for the truncated and full-length transcripts. Control experiments demonstrated that RNA preparation by the NP40 method, but not the GSCN/CsCl method, was efficient for the recovery of the truncated RNAs, which may account for the inconsistent observation of the truncated transcripts in HIV-1 transcriptional analysis.

What is the significance of the truncated transcripts? Activation of truncated transcripts may provide stringent control of HIV-1 gene expression in the absence of Tat. Differential expression of the cellular genes c-*myc*, c-*fos*, and c-*myb* in exponential and terminally differentiated cells is tightly controlled by programmed blocks to transcriptional elongation (Bentley and Groudine, 1986; Piecharczyk et al., 1987; Bender et al., 1987; Watson, 1987). Maintenance of high level basal transcription from the HIV-1 promoter may be important for the mechanism of Tat response.

The conclusion that sequences downstream of the RNA start site regulate HIV-1 basal transcription is not without precedent. Intragenic regions have been shown to mediate both repressive (bovine papilloma virus P7185 [Stenlund et al., 1990]) and activating effects (β-globin

gene [Charnay et al., 1984]; chicken thymidine kinase gene [Merrill et al., 1984]; adenovirus type 5 E1A promoter [Osbourne et al., 1984]; bovine leukemia virus promoter [Derse and Casey, 1986]; adenovirus major late promoter [Mansour et al., 1986]; c-*myc* P2 promoter [Yang et al., 1986]; Drosphilia *hsp*22 promoter [Hultmark et al., 1986]; human gastrin gene [Theill et al., 1987]; bovine papilloma virus P7185 [Stenlund et al., 1987]; minute virus of mice P38 [Krauskopf et al., 1990]). For example, in the adenovirus type 5 E1A promoter, a 10-fold reduction in steady-state RNA level is observed in response to a single base deletion 399 nucleotides downstream of the transcription start site (Osbourne et al., 1984). In the bovine leukemia virus (BLV) promoter, deletion of a positive control element in a 250 bp segment downstream of the RNA start site reduces gene expression by 87% (Derse and Casey, 1986). When this 250 bp position-dependent sequence was inserted downstream of the transcription start site of an SV40 early promoter, CAT activity increased 60-fold. In contrast to the upstream BLV enhancer, the 250 bp fragment is active in both BLV-infected and uninfected cells, suggesting a role for cellular factors (Derse and Casey, 1986). In the c-*myc* gene, a downstream transcription modulator is required for transcription from the c-*myc* P2 promoter (Yang et al., 1986). The modulator is orientation-dependent and promoter-specific, increasing transcription from the c-*myc* P2 promoter and SV40 upstream early promoter ($pSVE_2$) but not the c-*myc* P1 promoter or the SV40 downstream early promoter ($pSVE_1$; Yang et al., 1986). In minute virus of mice (MVM), a downstream 365 bp element (DPE) is essential for basal transcription from the P38 promoter (Krauskopf et al., 1990).

The basal transcription activity of the HIV-1 provirus may be one of many biological determinants of HIV-1 virulence (Golub et al., 1990; York-Higgins et al., 1990). Prior to synthesis of viral regulatory proteins, the rate of basal proviral transcription influences the subsequent viral replication cycle (Wong-Staal and Gallo, 1985; Tersemette et al., 1989). The experiments presented in this report indicate that the R region of the 5′ LTR plays an essential role in the activity of the basal Tat-independent HIV-1 promoter. Understanding the mechanism by which the R region is involved in regulation of HIV-1 transcription and the contributions of cellular proteins which control R region function may be fundamental to the design of therapies that delay productive HIV-1 infection.

### Acknowledgments

We thank Matt Rounsville for generous gifts of plasmids and Cindy Bohan for critical reading of the manuscript.

This work was funded in part by NIH grant AI25531 and AmFAR 000992-7RG in memory of Gerald S. Russell.

The costs of publishing this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

Kathleen A. Boris-Lawrie is currently at the McArdle Laboratory for Cancer Research, University of Wisconsin, 1400 University Ave., Madison, WI 53706.

## References

T. Barre-Sinoussi, J. C. Chermann, F. Reye, M. T. Nugeure, S. Charmaret, J. Gruest, C. Dauquet, C. Aixer-Blin, F. Vezinet-Brun, C. Rouzious, W. Rozenbaum, and L. Montagnier (1983), Science 220, 868–871.

A. Bender, C. B. Thomson, and W. M. Kuechl (1987), Science 237, 1473–1476.

D. L. Bentley and M. Groudine (1986), Nature 321, 702–706.

B. Berkhout and K.T. Jeang (1989), J Virol 63, 5501–5504.

B. Berkhout, R. H. Silverman, and K.T. Jeang (1989b), Cell 59, 273–282.

C. Bohan, R. A. Robinson, P. A. Luciw, and A. Srinivasan (1989), Virology 172, 573–583.

R. Bohnlein, J. W. Lowenthan, M. Siekevitz, D. W. Ballard, B. R. Franza, and W. C. Greene (1988), Cell 53, 827–836.

P. Charnay, R. Treisman, P. Mellon, M. Chao, R. Axel, and T. Maniatis (1984), Cell 38, 251–263.

J. J. Chirgwin, A. E. Przbyla, R. J. MacDonald, and W. J. Rutter (1979), Biochemistry 18, 5294–5299.

L. Clark, J. R. Matthews, and R. T. May (1990), J Virol 64, 1335–1344.

B. R. Cullen and W. C. Greene (1989), Cell 58, 423–426.

D. Derse and J. W. Casey (1986), Science 231, 1437–1440.

E. J. Duh, W. J. Maury, T. M. Folks, A. S. Fauci, and A. B. Rabson (1989), Proc Natl Acad Sci USA 86, 5974–5978.

A. S. Fauci (1988), Science 239, 617–622.

S. Feng and E. C. Holland (1988), Nature 334, 165–167.

T. M. Folks, J. Kelly, S. Benn, A. Kinter. J. Justement, J. Gold, R. Redfield, K. W. Sell, and A. S. Fauci (1986), J Immunol 136, 4049–4053.

B. R. Franza, Jr., S. F. Josephs, M. Z. Gilman, W. Ryan, and B. Clarkson (1987), Nature 330, 391–395.
R. C. Gallo, S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kapten, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, G. White, P. Foster, and P. D. Markham (1984), Science 224, 500–503.
J. A. Garcia, F. K. Wu, R. Mitsuyasu, and R. B. Gaynor (1987), EMBO J 6, 3761–3770.
J. A. Garcia, D. Marrich, E. Soultanakis, F. K. Wu, R. Mitsuyasu, and R. B. Gaynor (1989), EMBO J 8, 765–778.
H. E. Gendelman, W. Phelps, L. Feigenbaum, J. M. Ostrove, A. Adachi, P. Howley, G. Khoury, H. S. Ginsberg, and M. A. Martin (1986), Proc Natl Acad Sci USA 83, 9759–9763.
E. I. Golub, G. Li, and D. J. Volsky (1990), J Virol 64, 3654–3660.
M. E. Harper, L. M. Marselle, R. C. Gallo, and F. Wong-Staal (1986), Proc Natl Acad Sci USA 83, 772–776.
D. Harrich, J. Garcia, R. Mitsuyasu, and R. Gaynor (1990), EMBO J 9, 4417–4423.
J. Hauber and B. R. Cullen (1988), J Virol 62, 673–679.
D. D. Ho, T. Moudgil, and M. Alam (1989), N Engl J Med 321, 1621–1625.
D. Hultmark, R. Klemenz, and W. J. Gehring (1986), Cell 44, 429–438.
A. Jakobovits, D. H. Smith, E. B. Jakobovits, and D. J. Capon (1988), Mol Cell Biol 8, 2555–2561.
K. A. Jones, J. T. Kadonaga, P. A. Luciw, and R. Tjian (1986), Science 232, 755–759.
K. A. Jones, P. A. Luciw, and N. Duchange (1988), Genes Dev 2, 1101–1114.
S.-Y. Kao, A. F. Calman, P. A. Luciw, and B. M. Peterlin (1987), Nature 330, 489–493.
H. Kato, M. Horikoshi, and R. G. Roeder (1991), Science 251, 1476–1479.
K. Kawakami, C. Scheidereit, and R. Roeder (1988), Proc Natl Acad Sci USA 85, 4700–4704.
S. Kliewer, J. Garcia, L. Pearson, E. Soultanakis, A. Dasgupta, and R. B. Gaynor (1989), J Virol 63, 4616–4625.
A. Krauskopf, O. Resnekov, and Y. Aloni (1990), J Virol 64, 354–360.
M. F. Laspia, A. P. Rice, and M. B. Matthews (1989), Cell 59, 283–292.
J. Lee, Y. Li, S. Doerre, P. Sista, D. W. Ballard, W. C. Greene, and B. R. Franza (1991), Oncogene 6, 665–667.
J. Leonard, C. Parrott, A. J. Buckler-White, W. Turner, E. K. Ross, M. A. Martin, and A. B. Rabson (1989), J Virol 63, 4919–4924.
J. A. Levy, A. D. Hoffman, S. M. Kramer, J. A. Landis, J. M. Shimabukuro, and L. S. Oshino (1984), Science 225, 840–842.
M. H. Malim, R. Fenrick, D. W. Ballard, J. Hauber, E. Bohnlein, and B. R. Cullen (1989), J Virol 63, 3213–3219.
J. L. Manley, A. Fire, A. Cano, P. A. Sharp, and M. L. Gefter (1980), Proc Natl Acad Sci USA 77, 3855–3859.
S. L. Mansour, T. Grodzicker, and R. Tjian (1986), Mol Cell Biol 6, 2684–2694.
J. S. McDougal, A. Mawle, S. P. Cort, J. A. Nicholson, G. D. Cross, J. A. Scheppler-Campbell, D. Hicks, and J. Sligh (1985), J Immunol 135, 3151–3162.
G. F. Merrill, R. M. Harland, M. Groudine, and S. McKnight (1984), Mol Cell Biol 4, 1769–1776.
J. Muesing, D. H. Smith, and D. J. Capon (1987), Cell 48, 691–701.
G. Nabel and D. Baltimore (1987), Nature 326, 711–713.
T. F. Osbourne, D. N. Arvidson, E. S. Tyau, M. Dunsworth-Browne, and A. J. Berk (1984), Mol Cell Biol 4, 1293–1305.
M. Piechaczyk, J. M. Blanchard, and P. Jeanteur (1987), Trends Genet 3, 47–51.
R. Ratnasabapathy, M. Sheldon, L. Johal, and N. Hernandez (1990), Genes Dev 4, 2061–2074.
C. A. Rosen, J. G. Sodroski, and W. A. Haseltine (1985), Cell 41, 813–823.
S. Roy, N. T. Parkin, C. Rosen, J. Itovitch, and N. Sonenberg (1990), J Virol 64, 1402–1406.
R. M. Schmid, N. D. Perkins, C. S. Duckett, P. C. Andrews, and G. J. Nabel (1991), Nature 352, 733–736.
S. M. Schnittman, M. C. Psallidopoulos, H. C. Lane, L. Thompson, M. Baseler, F. Massari, C. H. Fox, N. P. Salzman, and A. S. Fauci (1989), Science 245, 305–308.
M. J. Selby, E. S. Baia, P. A. Luciw, and B. M. Peterlin (1989), Genes Dev 3, 547–558.
M. J. Selby and B. M. Peterlin (1990), Cell 62, 769–776.
J.-P. Shaw, P. J. Utz, D. B. Durband, J. J. Toole, E. A. Emmel, and G. R. Crabtree (1988), Science 241, 202–205.
M. Siekevitz, S. F. Josephs, M. Dukovich, N. Peffer, F. Wong-Staal, and W. C. Greene (1987), Science 238, 1575–1578.
C. Southgate, M. L. Zapp, and M. R. Green (1990), Nature 345, 640–642.
R. R. Spaete and E. S. Mocarski (1985), J Virol 56, 135–143.
A. Stenlund, G. L. Bream, and M. R. Botchan (1987), Science 236, 1666–1671.
A. Stenlund and M. R. Botchan (1990), Genes Dev 4, 123–136.
M. Tersemette, R. A. Gruters, F. de Wolf, R. E. Y. de Goede, J. M. A. Lange, P. T. A. Schellekens, J. Goudsmit, G. H. Huisman, and F. Miedema (1989), J Virol 63, 2118–2125.
L. E. Theill, O. Wiborg, and J. Vuust (1987), Mol Cell Biol 7, 4329–4336.
S. E. Tong-Starksen, P. A. Luciw, and B. M. Peterlin (1987), Proc Natl Acad Sci USA 84, 6845–6849.
M. G. Toohey and K. A. Jones (1989), Genes Dev 3, 265–282.

R. Watson (1987), Oncogene 2, 267–272.
F. Wong-Staal and R. C. Gallo (1985), Nature 317, 395–403.
C. M. Wright, B. K. Felber, H. Paskalis, and G. N. Pavlakis (1986), Science 234, 988–992.
F. K. Wu, J. A. Garcia, D. Harrich, and R. B. Gaynor (1988), EMBO J 7, 2117–2129.
J-Q. Yang, E. F. Remmers, and K. B. Marcu (1986), EMBO J 5, 3553–3562.
D. York-Higgins, C. Cheng-Mayer, D. Bauer, J. A. Levy, and D. Dino (1990), J Virol 64, 4016–4020.
J. A. Zack, A. J. Cann, J. P. Lugo, and I. S. Y. Chen (1988), Science 240, 1026–1029.
S. L. Zeichner, J. Y. H. Kim, and J. C. Alwine (1991), Gene Expr 1, 15–27.
M. Zucker (1987), Cold Spring Harbor Symp Quant Biol 52, 123.