The HIV-1 long terminal repeat contains an unusual element that induces the synthesis of short RNAs from various mRNA and snRNA promoters.

R Ratnasabapathy, M Sheldon, L Johal, et al.

Genes Dev. 1990 4: 2061-2074
Access the most recent version at doi:10.1101/gad.4.12a.2061

The HIV-1 long terminal repeat contains an unusual element that induces the synthesis of short RNAs from various mRNA and snRNA promoters

Ratneswaran Ratnasabapathy, Michael Sheldon,[1] Lynn Johal, and Nouria Hernandez

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724; [1]Genetics Program, State University of New York at Stony Brook, Stony Brook, New York 11794 USA

We describe an unusual element that activates the synthesis of short transcripts from a wide variety of mRNA and small nuclear RNA (snRNA) promoters, including the U6 RNA polymerase III promoter. This inducer of short transcripts (IST) is located between positions −5 and +82 relative to the cap site in the HIV-1 LTR. In the presence of IST, the total transcriptional activity of the different promoters is greatly increased, but the resulting additional RNA molecules are short, ending around position +60. IST is not the RNA target (TAR) for Tat *trans*-activation; however, because it relies entirely on cellular factors for activity, IST may serve to provide abundant RNA targets for Tat *trans*-activation without a requirement for full-length viral mRNA expression.

[*Key Words*: HIV-1; TAR element, Tat *trans*-activation; short transcripts]

Received August 28, 1990; revised version accepted September 28, 1990.

In several transcription units, including the SV40 and adenovirus 2 (Ad2) major late transcription units, the c-*myc* gene, the *Drosophila hsp70* gene, and the HIV-1 long terminal repeat (LTR), some transcription is confined under certain circumstances to the region proximal to the promoter (for review, see Spencer and Groudine 1990). In the cases where the products of such transcription are stable enough to be detectable in steady-state RNA, they are found to consist of short, nonpolyadenylated RNA molecules. Very little is known about the mechanisms that govern the formation of these short transcripts or about the role of the short transcripts themselves. In HIV-1, however, it is likely that the short transcripts provide the target for *trans*-activation by the viral protein Tat.

Tat is a potent *trans*-activator of HIV-1 gene expression and functions at least in part by increasing the amounts of HIV-1 mRNAs at the level of transcription (Hauber et al. 1987; Kao et al. 1987; Jakobovits et al. 1988; Rice and Mathews 1988; Sadaie et al. 1988). Tat affects both initiation and elongation of transcription (Laspia et al. 1989), although some assays detect only the effect on elongation (Kao et al. 1987). Tat is unusual in that it acts through an RNA element, the *trans*-activation response (TAR) element (Feng and Holland 1988; Berkhout et al. 1989; Selby et al. 1989; Roy et al. 1990; Southgate et al. 1990; Selby and Peterlin, 1990) encoded by sequences downstream of the HIV-1 mRNA initiation site, between nucleotides +18 and +44 (Rosen et al. 1985; Hauber and Cullen 1988; Jakobovits et al. 1988; Selby et al. 1989). The TAR element corresponds to the upper half of a bulged stem-and-loop structure that extends from nucleotides +1 to +59 (Muesing et al. 1987). The short transcripts encompass this structure, extending from nucleotide +1 to about +60, and are more abundant in the absence than in the presence of Tat (Kao et al. 1987).

How are the short transcripts generated? There is clearly no universal block to transcription around position +60, because deletion or insertion of the R region of the HIV-1 LTR into other transcription units does not have any effect on gene expression (Rosen et al. 1985; Cullen 1986; Peterlin et al. 1986; Wright et al. 1986; Muesing et al. 1987; Hauber and Cullen 1988). Instead, the HIV-1 promoter might direct the formation of transcription complexes unable to elongate transcription efficiently; the short transcripts would then be the remnants, stabilized by the binding of cellular proteins to the stem-and-loop structure (Gatignol et al. 1989; Gaynor et al. 1989; Marciniak et al. 1990), of longer randomly terminated RNA molecules. Indeed, in the human c-*myc* and perhaps the *Drosophila hsp70* genes, events occurring at the promoter can influence the elongation properties of the transcription complex (Bentley and Groudine 1988; Rougvie and Lis 1988; Spencer et al. 1990; for review, see Spencer and Groudine 1990).

Perhaps the best-characterized examples of interdependence between promoter and terminator sequences

are the human U1 and U2 small nuclear RNA (snRNA) genes, whose normal products are also short, nonpolyadenylated RNAs. U-snRNAs are involved in RNA processing (for review, see Steitz et al. 1988) and, except for U6, which is synthesized by RNA polymerase III from an unusual promoter located in the 5′-flanking sequence of the gene (Das et al. 1988), the snRNAs are synthesized by RNA polymerase II (for review, see Dahlberg and Lund 1988). The RNA polymerase II snRNA promoters direct the formation of specialized transcription complexes that, unlike mRNA transcription complexes, are capable of recognizing a termination signal called the 3′ box, located just downstream of the genes (for review, see Dahlberg and Lund 1988; Hernandez and Lucito 1988; Neuman de Vegvar and Dahlberg 1989). Furthermore, these specialized snRNA transcription complexes are not capable of efficient elongation, because in the absence of the 3′ box they terminate heterogeneously at nearby downstream cryptic 3′ boxes that sometimes bear little resemblance to the authentic signal (Neuman de Vegvar et al. 1986). Thus, here, the fate of a transcription complex with regard to elongation and termination is determined at the promoter.

To determine whether formation of short transcripts in the HIV-1 transcription unit is also dependent on promoter sequences, we have tested a series of hybrid constructs in which the HIV-1 promoter sequences are replaced by other mRNA and snRNA promoters. These experiments have led us to the discovery of a novel element, located between positions −5 and +82 relative to the transcription initiation site in the HIV-1 LTR, that activates the synthesis of short transcripts from every promoter we have tested including the U2 and the RNA polymerase III U6 promoters. We call this element IST, for inducer of short transcripts. In our assay, IST does not interfere significantly with the basal production of full-length transcripts from the various promoters, but instead superimposes itself on the normal promoter activity to make additional short transcripts. Hence, mRNA synthesis is not increased significantly although the promoter is much more active. Mutations that inactivate TAR do not affect IST function, showing that these are two different elements. Because IST relies entirely on cellular factors for function, it is likely to have cellular counterparts and may be an example of a new category of regulatory elements that affect transcription both at the level of initiation and of elongation.

Results

Experimental design

To examine directly the amounts of short and long RNAs produced from the HIV-1 promoter, we designed a transcription unit short enough so that both short and long transcripts could be easily probed over their entire length by RNase T_1 mapping. This "mini" transcription unit, pHIV-1/R, is shown in Figure 1A. It contains HIV-1 sequences from −138 to +82 relative to the cap site (+1), and thus comprises the HIV-1 promoter region as well as nearly the entire R region of the HIV-1 LTR, including the sequences encoding the TAR RNA element. Immediately downstream of the R region is a fragment derived from Ad2 that contains the L3 polyadenylation signal. Farther downstream, we inserted an antisense bacteriophage T3 promoter such that perfectly complementary riboprobes could be synthesized easily from this and all subsequent constructs described below. pHIV-1/R also carries the SV40 origin of replication, so that its copy number can be increased where necessary by coexpression of SV40 T antigen (T Ag).

To characterize the pHIV-1/R transcription unit, this vector was transfected into Jurkat T cells, a host for HIV-1 replication, together with a construct expressing the Tat protein or a similar construct lacking the Tat-coding sequences. A plasmid containing the human α-globin gene, a gene not responsive to Tat, was included as an internal reference in this and all subsequent transfections. RNA extracted from the transfected cells was analyzed by RNase T_1 protection mapping with a mixture of two probes—probe α98, which is protected over 98 nucleotides by correctly initiated α-globin RNA, and either probe T3/534 or probe T3/358 (Fig. 1A). Probe T3/534 extends from downstream of the polyadenylation site to upstream of the start site of transcription and therefore simultaneously maps the 5′ and 3′ ends of correctly initiated and polyadenylated RNA molecules (Fig. 1A). Probe T3/358 is truncated at position +22 downstream of the HIV-1 transcriptional start site and should, therefore, give rise to fragments 22 nucleotides shorter than probe T3/534 upon protection by correctly initiated HIV-1 RNA (Fig. 1A).

The patterns obtained in such an analysis are shown in Figure 1B. The constant intensity of the α-globin signal (labeled α) shows that the transfection efficiencies and RNA recoveries were very similar in each pair of samples. The bands labeled RT, L3, and st all correspond to RNA correctly initiated at the HIV-1 cap site, because the RNAs were shortened by the expected length when the T3/358 probe was used (cf. lanes 1 and 2 with lanes 3 and 4). The lengths of these different protected fragments indicate that the band labeled RT corresponds to RNA reading through the L3 polyadenylation site, L3 corresponds to RNA polyadenylated at the L3 site, and the bands labeled st correspond to short RNAs ending heterogeneously downstream of the stem-and-loop structure, around position +60. Fractionation of the RNAs over oligo(dT)-cellulose columns confirmed that the L3 RNA is polyadenylated, whereas the RT and st transcripts are not polyadenylated (data not shown). The bands labeled st seemed, then, to correspond to the short transcripts observed by Kao et al. (1987); however, these bands could also result from artifactual RNase T_1 cleavage of the probe in the RNA–RNA hybrids around position +60. To exclude this possibility, we replaced the HIV-1 promoter in pHIV-1/R by the bacteriophage T7 promoter and synthesized unlabeled RNA extending through the whole transcription unit with bacteriophage T7 RNA polymerase. This RNA was then hybridized to

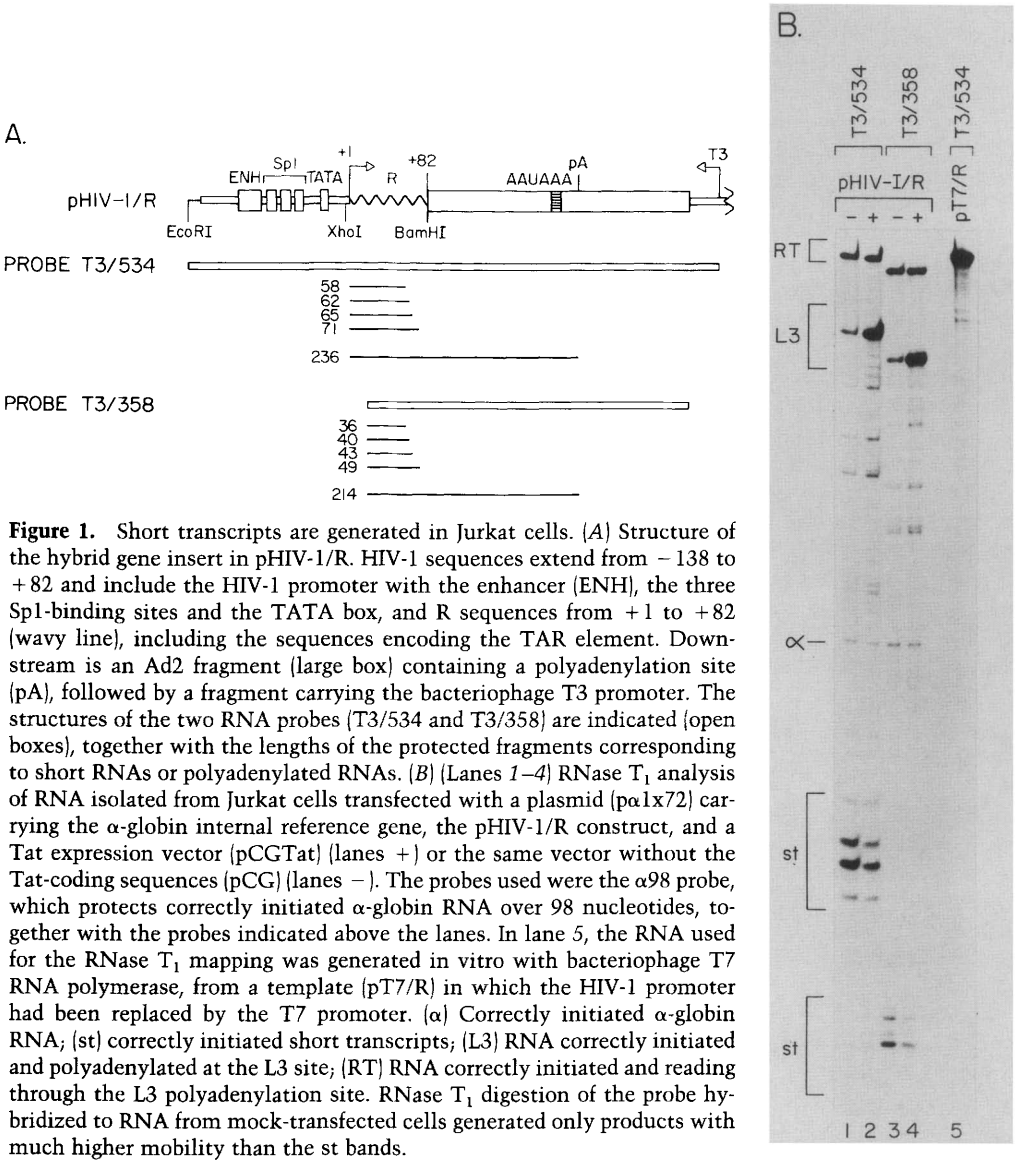


Figure 1. Short transcripts are generated in Jurkat cells. (*A*) Structure of the hybrid gene insert in pHIV-1/R. HIV-1 sequences extend from −138 to +82 and include the HIV-1 promoter with the enhancer (ENH), the three Sp1-binding sites and the TATA box, and R sequences from +1 to +82 (wavy line), including the sequences encoding the TAR element. Downstream is an Ad2 fragment (large box) containing a polyadenylation site (pA), followed by a fragment carrying the bacteriophage T3 promoter. The structures of the two RNA probes (T3/534 and T3/358) are indicated (open boxes), together with the lengths of the protected fragments corresponding to short RNAs or polyadenylated RNAs. (*B*) (Lanes *1–4*) RNase T_1 analysis of RNA isolated from Jurkat cells transfected with a plasmid (pα1x72) carrying the α-globin internal reference gene, the pHIV-1/R construct, and a Tat expression vector (pCGTat) (lanes +) or the same vector without the Tat-coding sequences (pCG) (lanes −). The probes used were the α98 probe, which protects correctly initiated α-globin RNA over 98 nucleotides, together with the probes indicated above the lanes. In lane *5*, the RNA used for the RNase T_1 mapping was generated in vitro with bacteriophage T7 RNA polymerase, from a template (pT7/R) in which the HIV-1 promoter had been replaced by the T7 promoter. (α) Correctly initiated α-globin RNA; (st) correctly initiated short transcripts; (L3) RNA correctly initiated and polyadenylated at the L3 site; (RT) RNA correctly initiated and reading through the L3 polyadenylation site. RNase T_1 digestion of the probe hybridized to RNA from mock-transfected cells generated only products with much higher mobility than the st bands.

the antisense probe T3/534, and the hybrids were treated with RNase T_1 under the same conditions as the samples in lanes 1 and 2. As shown in lane 5, RNase T_1 mapping of this RNA with probe T3/534 protected the entire RNA molecule and did not generate any short fragments, demonstrating that the short fragments detected with RNA from transfected cells correspond to molecules ending around position +60.

In the presence of Tat (Fig. 1B, lanes 2 and 4), there was a slight decrease in the amounts of short transcripts, together with an increase in the amounts of polyadenylated RNA. Surprisingly, the amounts of correctly initiated RNA reading through the polyadenylation site (RT transcripts) were not increased in the presence of Tat. This observation is reproducible and may indicate that these transcripts are synthesized aberrantly in a nuclear compartment devoid of polyadenylation and Tat *trans*-activation machinery. Nevertheless, the experiment in Figure 1 shows that the pHIV/R transcription unit can be used to map simultaneously short and long transcripts in their entirety and that it responds to Tat *trans*-activation. The patterns of short and long transcripts in the presence and absence of Tat in Jurkat T cells are consistent with the results of Kao et al. (1987) in COS cells.

To determine the role of the HIV-1 promoter sequences in the formation of short transcripts, we constructed several derivatives of the parent pHIV-1/R construct in which the HIV-1 promoter was exchanged for other mRNA promoters, for the U2 snRNA promoter, which is known to direct the formation of a transcription complex with different elongation properties, and for the RNA polymerase III U6 promoter. RNA expression from the different constructs was monitored by RNase T_1 mapping and run-on experiments. The constructs were tested, whenever possible, in Jurkat cells, a

T-cell line, because T cells are a target for HIV-1 infection. Some promoters were not expressed efficiently in Jurkat cells, however, and were therefore tested in HeLa cells, in some cases in the presence of SV40 T Ag to amplify the signal. The run-on experiments were all performed in COS cells, because amplification of the templates was more efficient than in HeLa cells cotransfected with a T Ag expression vector. However, the results in all the cell lines were qualitatively identical.

Below, we first show that formation of short transcripts is not dependent on the HIV-1 promoter, but instead that the HIV-1 sequences between −5 and +82 induce the formation of short transcripts when located downstream of different mRNA promoters, the U2 promoter, and even the U6 promoter. We then show by run-on analyses that induction of the short transcripts correlates with an increase in transcription.

Short transcripts are synthesized from all tested mRNA promoters fused to HIV-1 sequences from positions −5 to +82

To determine whether the formation of short transcripts is mRNA promoter dependent, we tested constructs in which the HIV-1 promoter was replaced by the β-globin promoter (pβgl/R), the Ad2 major late promoter (pMLP/R), and the α-globin promoter (pαgl/R) (Fig. 2A). Figure 2B shows the patterns obtained after transfection into HeLa cells. As observed in Jurkat cells, the wild-type construct, pHIV-1/R, generated mainly short transcripts in the absence of Tat (lane 1). In the presence of Tat, the amounts of short transcripts were decreased significantly and polyadenylated RNA was increased ~40-fold, as determined by scintillation counting of excised bands (lane 2). A 4-bp deletion designated +34Δ4, shown pre-

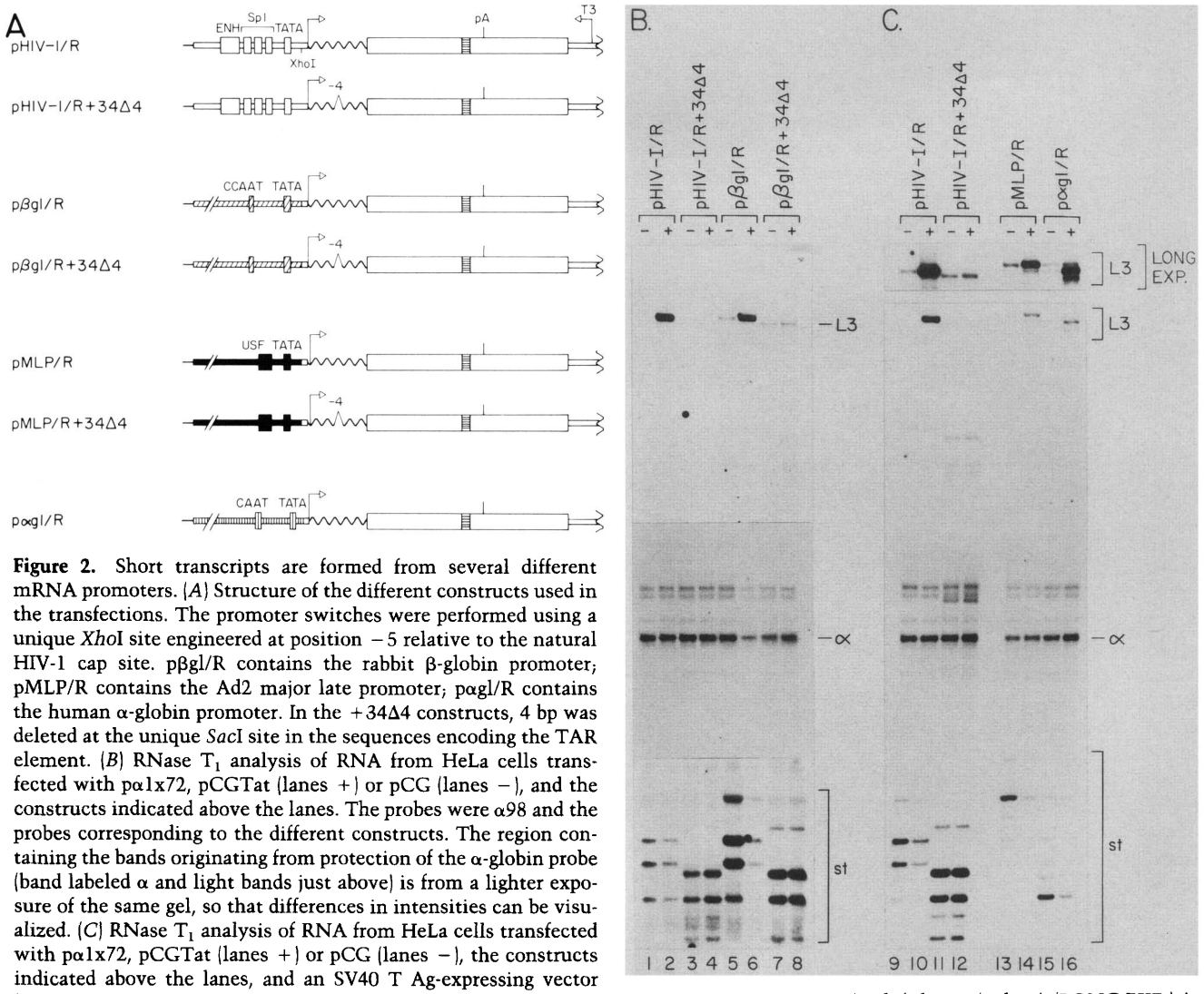


Figure 2. Short transcripts are formed from several different mRNA promoters. (*A*) Structure of the different constructs used in the transfections. The promoter switches were performed using a unique *Xho*I site engineered at position −5 relative to the natural HIV-1 cap site. pβgl/R contains the rabbit β-globin promoter; pMLP/R contains the Ad2 major late promoter; pαgl/R contains the human α-globin promoter. In the +34Δ4 constructs, 4 bp was deleted at the unique *Sac*I site in the sequences encoding the TAR element. (*B*) RNase T_1 analysis of RNA from HeLa cells transfected with pα1x72, pCGTat (lanes +) or pCG (lanes −), and the constructs indicated above the lanes. The probes were α98 and the probes corresponding to the different constructs. The region containing the bands originating from protection of the α-globin probe (band labeled α and light bands just above) is from a lighter exposure of the same gel, so that differences in intensities can be visualized. (*C*) RNase T_1 analysis of RNA from HeLa cells transfected with pα1x72, pCGTat (lanes +) or pCG (lanes −), the constructs indicated above the lanes, and an SV40 T Ag-expressing vector (pSVEori-; Hernandez and Lucito 1988). The probes were as in *B*. *C* is a much shorter exposure (10 hr) than *B* (5 days). (LONG EXP.) A long exposure of the region containing the L3 bands. Note that the sizes of the short transcripts vary depending on the exact site of initiation and the location of G residues in the probes, since RNase T_1 cleaves only after G residues. In addition, in the +34Δ4 constructs, each RNA species is shortened by 4 nucleotides, the extent of the deletion.

viously to debilitate the TAR element (deletion pD+35/+38 in Hauber and Cullen, 1988), abolished both the Tat-directed increase in long transcripts and the decrease in short transcripts (lanes 3 and 4), demonstrating that these effects are TAR dependent. The +34Δ4 deletion did not, however, abolish the formation of short transcripts (Fig. 2B, lanes 3 and 4; note that the short transcripts are shortened by the extent of the deletion). Thus, as described previously (Selby et al. 1989), formation of the short transcripts is not dependent on a functional TAR element.

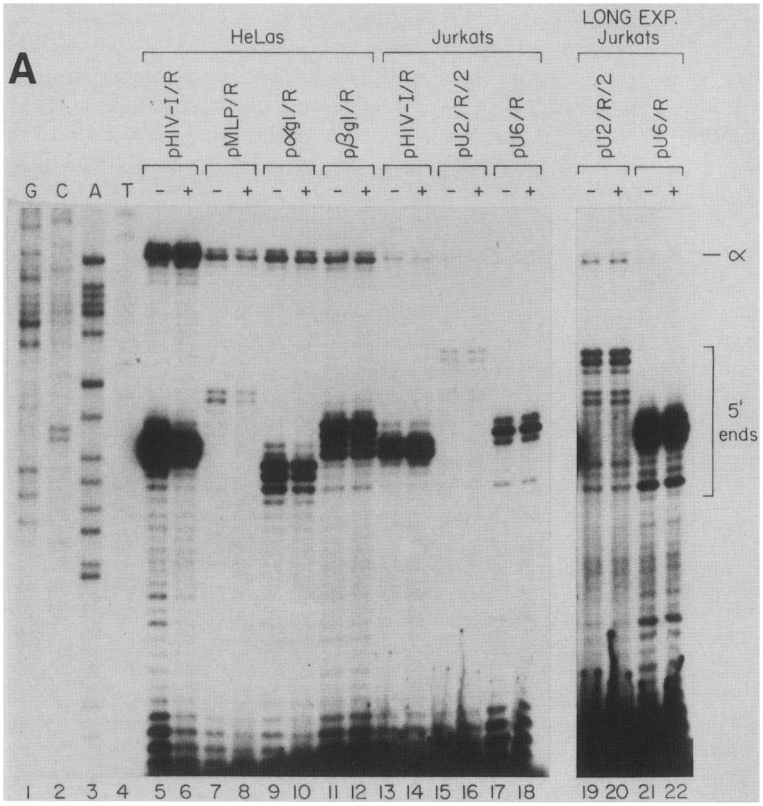
Replacement of the HIV-1 promoter by the β-globin promoter had no deleterious effect on either formation of short transcripts or Tat *trans*-activation (Fig. 2B, lanes 5 and 6; note that the sample in lane 6 is under-represented, as evidenced by the low α-globin signal). As observed with the HIV-1 promoter, Tat diminished the amounts of short transcripts and increased the amounts of polyadenylated RNA (20-fold), and this effect was dependent on TAR, but formation of the short transcripts itself was not dependent on an intact TAR element (lanes 7 and 8). The α-globin and Ad2 major late promoter fusions also gave similar results (Fig. 2C, lanes 13–16), even though SV40 T Ag was used to amplify the signal [note that T Ag did not affect the pattern of transcription from the HIV-1 promoter, except to amplify the signal (Fig. 2C, lanes 9–12)]. The levels of Tat *trans*-activation, however, were lower in these two cases (about fivefold). These results show that the formation of short transcripts is not dependent on HIV-1 promoter sequences upstream of position −5, and, therefore, the short transcripts are unlikely to arise because the HIV-1 upstream promoter sequences establish a transcription complex incapable of efficient elongation as is the case in the U2 snRNA promoter.

The observation that coexpression of Tat both increases the amounts of polyadenylated RNA and decreases the amounts of short transcripts prompted us to investigate whether in our assay system Tat increased the total levels of RNAs or only affected the proportion of short and long transcripts derived from the promoter. For this purpose, we quantitated the bands corresponding to short and long transcripts by liquid scintillation counting. Surprisingly, we found that the molar decrease in short transcripts was either larger or equivalent to the molar increase in long transcripts (data not shown). It is often difficult to observe this correspondence visually because the short transcripts are frequently in a greater than 50-fold molar excess over the long transcripts. Thus, only a 20% decrease in the short transcripts can account for a 10-fold increase in the less abundant full-length transcripts. To assay directly for changes in total RNA levels, we performed primer extension analyses using a primer (designated tarloop primer) that anneals to sequences between +32 and +59 and therefore scores both short and long transcripts. As an internal control, an α-globin RNA-specific primer was included. Figure 3A shows the results of analysis of RNA from both HeLa (Fig. 2B,C) and Jurkat transfections. In most of the constructs, the transcription initiation sites were displaced compared to the initiation site in the HIV-1 R construct. The start sites corresponding to the longest primer extension product for each construct are indicated by a star in Figure 3B. They were consistently located close to the position (underlined nucleotides) of the start sites in the corresponding wild-type promoters, suggesting that their location is directed by the upstream promoter elements. The striking observation was, however, that for each of the constructs the levels of total correctly initiated RNA in the presence of Tat were equal to or lower than the levels in the absence of Tat. Thus, in our assay, Tat did not increase initiation of transcription. These observations are consistent with the results of Kao et al. (1987), but do not themselves establish the mechanism of Tat *trans*-activation (see Discussion).

Together, the experiments above demonstrate that formation of short transcripts and Tat *trans*-activation are not dependent on HIV-1 promoter sequences upstream of −5, but instead can be observed with several different mRNA promoters. Furthermore, formation of the short transcripts is not dependent on an intact TAR element.

A U2 snRNA promoter–R construct directs the formation of short transcripts and is trans*-activated by Tat*

To assess the generality of short transcript formation and Tat *trans*-activation, we replaced the HIV-1 promoter by the U2 promoter, which consists of a snRNA-specific proximal element that is centered around −50 and is functionally equivalent to the TATA box of mRNA promoters, and an enhancer composed of an octamer motif adjacent to an Sp1-binding site (for review, see Dahlberg and Lund 1988). In addition, we inserted a fragment containing the last 6 nucleotides of the U1-coding sequence and the U1 3′ box before the Ad2 polyadenylation site (constructs pU2/R/2 and pU2/R+34Δ4/2; Fig. 4A). The resulting constructs contain two consecutive 3′-end formation signals downstream of the HIV-1 sequences: the U1 3′ box and the L3 polyadenylation site. Insertion of the U1 3′ box into the parent constructs pHIV-1/R and pHIV-1/R+34Δ4 (see Fig. 4A, constructs pHIV-1/R/2 and pHIV-1/R+34Δ4/2) did not have any effect on the pattern of RNA molecules generated from the HIV-1 promoter in the presence and absence of Tat (Fig. 4B, lanes 1–4), confirming that transcription complexes initiated at mRNA promoters ignore the snRNA 3′ box termination signal (Hernandez and Weiner 1986; Neuman de Vegvar et al. 1986). In contrast, the U2 promoter gave rise to new RNA species terminating at the U1 3′ box, as evidenced by the appearance of a typical doublet of bands in the RNase T_1 protection pattern (lanes 5–8, bands labeled U1 3′ end). As expected, in the absence of Tat, very little RNA derived from the U2 promoter extended through the 3′ box to the polyadenylation site (lane 5, band labeled L3). But unexpectedly, this construct gave rise to short transcripts (lane 5, bands labeled st), suggesting that the element that directs the formation of short transcripts in the HIV-1 −5 to +82



B

	-5 Xho I ↓ *
pHIV-1/R	AGCTGCTTTTc tCga GTACTGGGTCTCTCT
pMLP/R	GGGGGCGCtag .... ..*_..........
pαgl/R	ACCCTGGCGCG ....*._.....
pβgl/R (Fig. 2)	GAGCAGGGCAG ....*_.....
pU2/R	GCGAAGGCccc ..*_. .....
pU6/R	TGTGGAAAccc ....*_.....

Figure 3. Primer extension analysis of RNA derived from the different constructs. (*A*) Primer extension analysis of RNA from HeLa cells (lanes *5–12*) transfected with pα1x72, pCGTat (lanes +) or pCG (lanes −), the constructs indicated above the lanes, and the SV40 T Ag-expressing vector (pSVEori-), or from Jurkat cells (lanes *13–22*) transfected with the same constructs but in the absence of the SV40 T Ag-expressing vector (lanes *13–22*). The primers were the tarloop primer and a primer complementary to α-globin RNA. (Lanes *1–4*) Sequence ladders obtained by extension of the tarloop primer on the pHIV-1/R template in the presence of dideoxynucleotides. (Lanes *19–22*) Long exposure of lanes *15–18*). (*B*) Sequence around the transcription start site in the indicated constructs. The sequences upstream of the *Xho*I site are derived from the different promoters indicated in the constructs' names; the sequences downstream of the *Xho*I site are HIV-1 sequences. Lowercase characters indicate nucleotides that differ from the wild-type sequence of the different promoters. Stars indicate the start sites corresponding to the longest primer extension products in *A*; underlined nucleotides indicate the positions of the start sites in the wild-type promoters.

sequences can act on transcription complexes derived from a wide variety of RNA polymerase II promoters.

When the Tat protein was coexpressed with the U2 promoter construct, the levels of short transcripts decreased, but the levels of RNAs ending at the 3′ box did not increase; in fact, they were also diminished (Fig. 4B, lane 6). In contrast, the levels of polyadenylated RNAs increased significantly (10-fold, cf. lanes 6 and 5). This effect is TAR dependent, since it was abolished by the +34Δ4 mutation (lanes 7 and 8). This is a surprising result because it suggests that Tat can modify the elongation properties of a U2 transcription complex such that it can read through the 3′ box termination signal, like an mRNA transcription complex. To exclude the separate possibility that Tat was *trans*-activating a cryptic, mRNA-type promoter overlapping with the U2 promoter, we tested a construct in which the snRNA-specific proximal promoter element was debilitated by a clustered point mutation (pU2PE$^-$/R/2; Fig. 4A). This construct produced no detectable RNA in the absence and presence of Tat (lanes 9 and 10), indicating that the effect of Tat seen with the wild-type U2 construct occurs through the snRNA promoter. Thus, the HIV-1 sequences from −5 to +82 can direct the formation of short transcripts from a very unusual RNA polymerase II promoter, and Tat can change the elongation properties of the U2 transcription complex such that it becomes similar to a mRNA complex and capable of reading through a 3′ box. At present we do not know where this change in elongation properties occurs (e.g., before or after initiation of transcription).

A U6 promoter–R construct directs the formation of short transcripts but is not trans*-activated by Tat*

The RNA polymerase III U6 promoter is nearly identical in structure to the snRNA polymerase II promoters except for the presence of a third element centered around −25, the A/T box, which determines the RNA polymerase III specificity of the promoter (Mattaj et al. 1988; Lobo and Hernandez 1989, and references therein). Since the essential promoter elements of the U6 gene lie 5′ of the transcriptional start site, we could replace the HIV-1 promoter with the U6 promoter without affecting the −5 to +82 HIV-1 sequences (construct pU6/R, Fig. 5A) and then test whether even an RNA polymerase III promoter can direct formation of short transcripts and be *trans*-activated by Tat. RNA polymerase III transcription is terminated by stretches of T residues; such a run of T residues is provided in pU6/R within the Ad2 fragment, before the polyadenylation signal (see Fig. 5A). As shown in Figure 5B, lane 3, no major correctly initiated RNA species produced by this construct extended past the run of T residues, as expected. However, a large amount of short transcripts was also generated; these short transcripts were derived from the U6 promoter because mutation of the U6 proximal element resulted in a

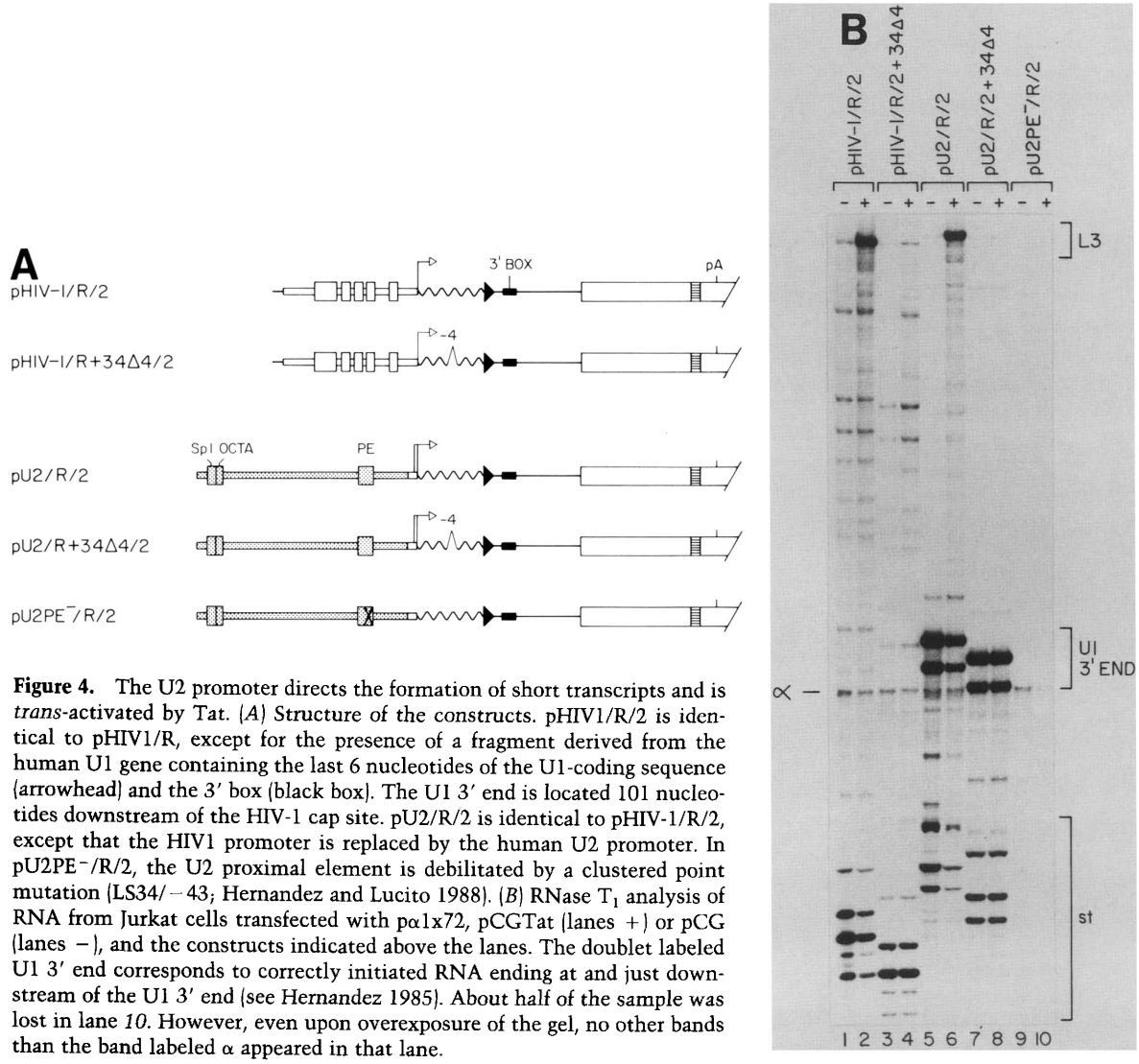


Figure 4. The U2 promoter directs the formation of short transcripts and is *trans*-activated by Tat. (*A*) Structure of the constructs. pHIV1/R/2 is identical to pHIV1/R, except for the presence of a fragment derived from the human U1 gene containing the last 6 nucleotides of the U1-coding sequence (arrowhead) and the 3′ box (black box). The U1 3′ end is located 101 nucleotides downstream of the HIV-1 cap site. pU2/R/2 is identical to pHIV-1/R/2, except that the HIV1 promoter is replaced by the human U2 promoter. In pU2PE⁻/R/2, the U2 proximal element is debilitated by a clustered point mutation (LS34/−43; Hernandez and Lucito 1988). (*B*) RNase T_1 analysis of RNA from Jurkat cells transfected with pα1x72, pCGTat (lanes +) or pCG (lanes −), and the constructs indicated above the lanes. The doublet labeled U1 3′ end corresponds to correctly initiated RNA ending at and just downstream of the U1 3′ end (see Hernandez 1985). About half of the sample was lost in lane *10*. However, even upon overexposure of the gel, no other bands than the band labeled α appeared in that lane.

loss of both full-length and short transcripts (lanes 5 and 6). In sharp contrast to the RNA polymerase II promoter constructs, however, Tat had no effect on the ratio of long and short transcripts (Fig. 5B, cf. lanes 3 and 4 with the Tat *trans*-activation of the HIV-1 promoter construct in lanes 1 and 2), indicating that Tat is not able to affect transcription derived from an RNA polymerase III promoter. These results show that formation of short transcripts can be uncoupled from Tat *trans*-activation in at least two ways: by replacement of the HIV-1 promoter with an RNA polymerase III promoter, as shown in Figure 5, and by inactivation of the TAR element with the +34Δ4 mutation, as shown previously (Selby et al. 1989) and in Figures 2 and 4 (and Fig. 6, below).

The HIV-1 sequences from −5 to +82 promote an increase in steady-state RNA levels derived from different promoters

The data described above show that the HIV-1 promoter region can be replaced by other RNA polymerase II promoters without abolishing the formation of short transcripts or Tat *trans*-activation. They are consistent with the HIV-1 LTR sequences from −5 to +82 directing 3′-end formation around position +60, except that they fail to explain why deletion of such a putative terminator does not result in an increase of basal gene expression (Rosen et al. 1985; Cullen 1986; Peterlin et al. 1986; Wright et al. 1986; Muesing et al. 1987; Hauber and Cullen 1988). To resolve this paradox, we replaced the HIV-1 sequences from −5 to +82 in the pHIV-1/R and pβgl/R constructs by the β-globin sequences naturally found from −5 to +82 in the β-globin gene. The resulting constructs contain either the HIV-1 promoter or the β-globin promoter fused to β-globin sequences, followed by the Ad2 fragment containing the L3 polyadenylation site (Fig. 6A). Thus, the pβgl/βgl construct does not contain any HIV-1 sequences.

Figure 6B compares the RNAs produced by the constructs containing the −5 to +82 HIV-1 sequences (pHIV-1/R, pHIV-1/R+34Δ4, pβgl/R, and pβgl/R+34Δ4; shown above in Fig. 2B) with the constructs containing

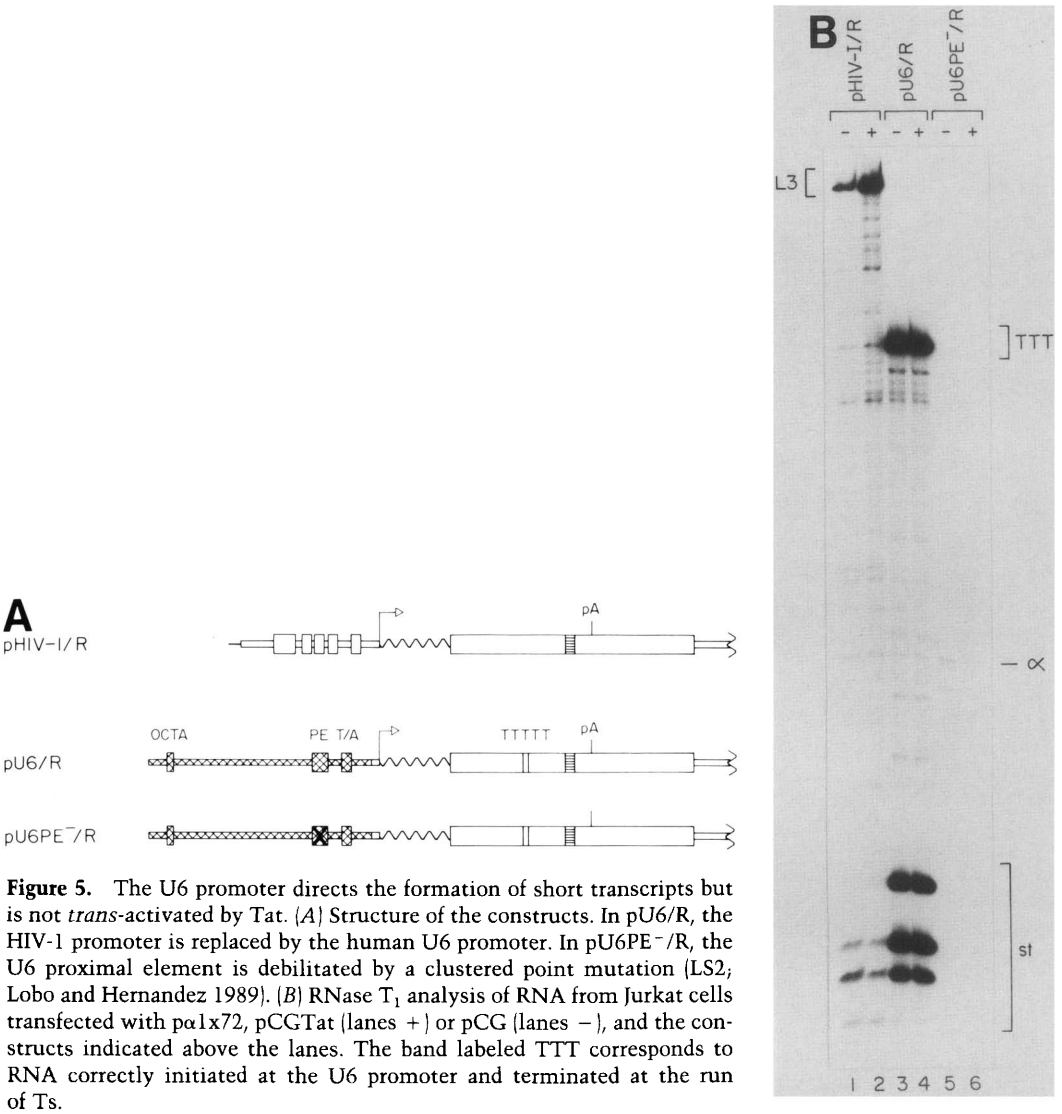


Figure 5. The U6 promoter directs the formation of short transcripts but is not *trans*-activated by Tat. (*A*) Structure of the constructs. In pU6/R, the HIV-1 promoter is replaced by the human U6 promoter. In pU6PE⁻/R, the U6 proximal element is debilitated by a clustered point mutation (LS2; Lobo and Hernandez 1989). (*B*) RNase T_1 analysis of RNA from Jurkat cells transfected with pα1x72, pCGTat (lanes +) or pCG (lanes −), and the constructs indicated above the lanes. The band labeled TTT corresponds to RNA correctly initiated at the U6 promoter and terminated at the run of Ts.

the −5 to +82 β-globin sequences (pHIV-1/βgl and pβgl/βgl) in the presence and absence of Tat. Surprisingly, the constructs containing the −5 to +82 β-globin sequences (lanes 6 and 7 and 12 and 13) did not direct formation of short transcripts and the levels of polyadenylated RNA were low; indeed, they were slighty lower than those produced by the constructs containing the −5 to +82 HIV-1 sequences in the absence of Tat (cf. lane 6 to lanes 2 and 4 and lane 12 to lanes 8 and 10, L3 band). Thus, the total number of RNA molecules derived from the pHIV1/βgl and pβgl/βgl constructs were much lower than the total number of RNA molecules, short and long, derived from the pHIV-1/R, pHIV/R+34Δ4, pβgl/R, and pβgl/R+34Δ4 constructs. Similar results were observed when the U6 promoter replaced the HIV-1 promoter in pHIV-1/R and pHIV-1/βgl (data not shown).

These results suggest that the HIV-1 region from −5 to +82 not only directs formation of the short transcripts but also activates the very transcription that results in their synthesis. However, an alternative interpretation was that the transcripts derived from the constructs containing the −5 to +82 β-globin sequences were much less stable than those derived from the constructs containing the −5 to +82 HIV-1 sequences, perhaps because they could not fold into a stable stem-and-loop structure. To address this question, we replaced the HIV-1 sequence from position +2 to +59 in pHIV-1/R with a synthetic DNA fragment corresponding to the +2 to +59 sequence read in a 3′ to 5′ direction (+59, +58, +57, etc.), to generate the clone pHIV-1/RAT. As indicated in Figure 6C, the RNA encoded by this sequence can potentially fold into a secondary structure of very similar stability as the stem-and-loop structure assumed by HIV-1 RNA. However, like pHIV-1/βgl, pHIV-1/RAT failed to produce detectable amounts of short transcripts and produced low levels of polyadenylated RNA (Fig. 6B, lane 1), sug-

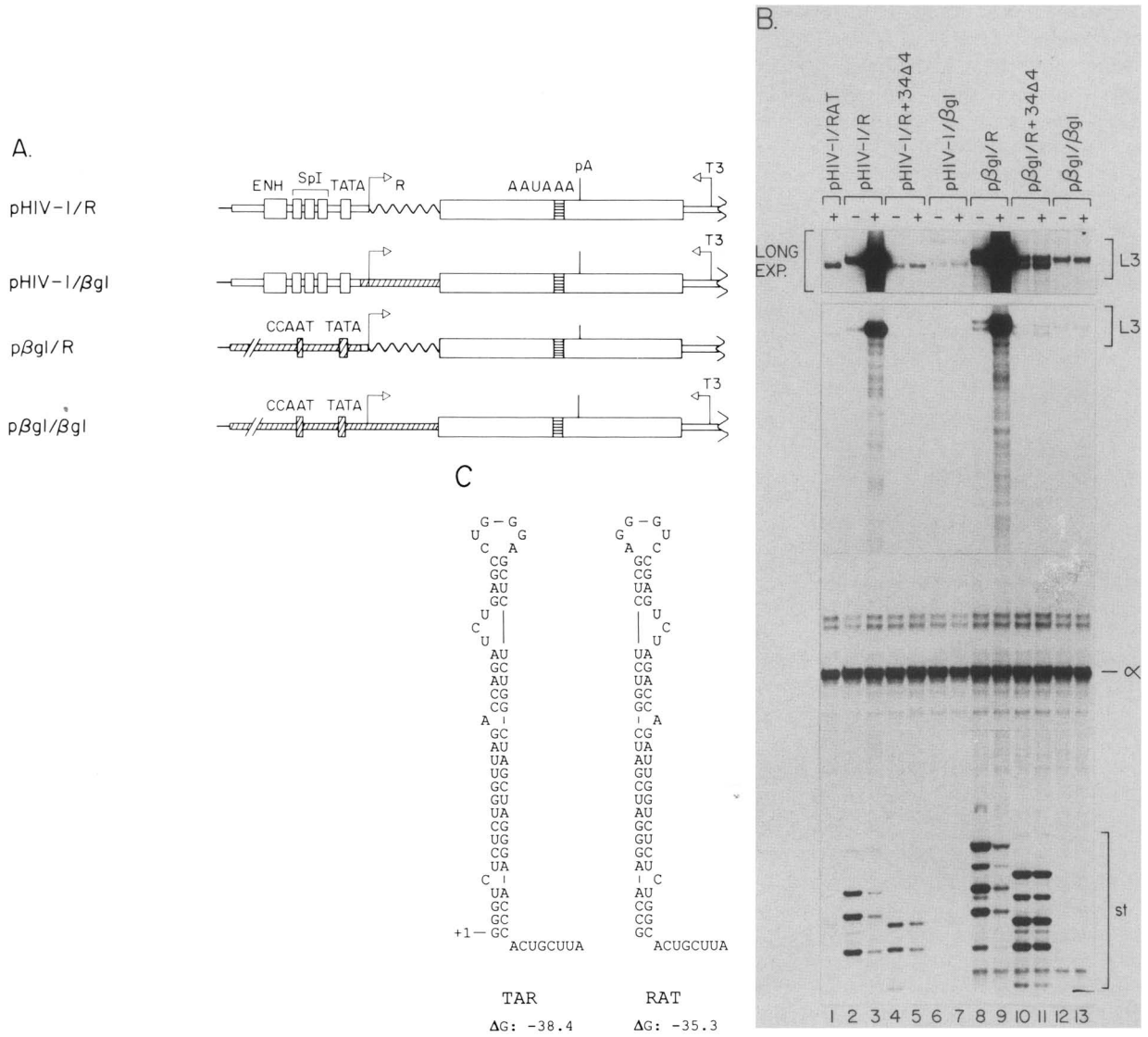


Figure 6. HIV-1 sequences from −5 to +82 increase the amounts of steady-state RNA derived from different promoters. (*A*) Structure of the constructs. pHIV-1/βgl is similar to pHIV-1/R, and pβgl/βgl is similar to pβgl/R except that the −5 to +82 HIV-1 sequences in these two constructs were replaced by rabbit β-globin sequences from −5 to +82. As a result, pβgl/βgl does not contain any HIV-1 sequences. (*B*) RNase T_1 analysis of RNA from HeLa cells transfected with pα1x72, pCGTat (lanes +) or pCG (lanes −), and the constructs indicated above the lanes. (LONG EXP.) A long exposure of the region containing the L3 bands. The region containing the bands originating from protection of the α-globin probe (band labeled α and light bands just above) is from a lighter exposure of the same gel. The sequence of the pβgl/R clone used in this experiment has 1 nucleotide inserted at the junction between the β-globin and HIV-1 sequences as compared to the pβgl/R clone used in Fig. 2 (see Materials and methods); this results in transcription initiating at two positions; hence, the doublet L3 bands in lanes *8–11*. (*C*) Predicted stem-and-loop structure in the pHIV-1/RAT construct. The free-energy values were calculated with the PC/GENE RNAFOLD program, which uses the method of Zucker and Stiegler (1981). The stability assumed by RNA derived from pHIV-1/RAT is very close to that of the stem-and-loop structure assumed by the 5′ region of HIV-1 RNAs (Muesing et al. 1987).

gesting that the large amounts of short transcripts observed in the presence of the −5 to +82 HIV-1 sequences do not arise solely because of folding of the RNA into a stable stem-and-loop structure. Thus, a large decrease in the total amount of RNA was observed with two substitutions of the −5 to +82 HIV-1 sequences (βgl and RAT), one of them encoding a potential RNA secondary structure very similar to that encoded by the wild-type HIV-1 R region.

The HIV-1 sequences from −5 to +82 stimulate transcription from RNA polymerase II and III promoters in a run-on assay

To determine directly whether the −5 to +82 HIV-1 region increases the steady-state levels of short RNAs by increasing the rate of transcription, we performed "run-on" experiments. Such experiments measure the amounts of nascent RNA molecules produced on a given

portion of the DNA template and thus reflect the density of RNA polymerases in that template region. COS cells were transfected with the reference α-globin gene and test constructs either containing the −5 to +82 HIV-1 region or with these sequences substituted by β-globin sequences. Twenty-four hours after transfection, the cells were permeabilized and exposed for 6 min to [α-^{32}P]UTP. The resulting radioactive RNA was hybridized to single-stranded probes immobilized on filters that were directed to the sense (ss+) or antisense (ss−) strands of the −5 to +82 HIV-1 region (probe mp18/R) or the −5 to +82 β-globin region (probe mp18/βgl). Probes that hybridize to the sense and antisense strands of the α-globin reference gene (probes mp18/α ss+ and ss−) were also included. The filters were then treated with RNase A, and the resistant signal was visualized by autoradiography.

Figure 7 shows the results of such an experiment for two pairs of constructs containing an RNA polymerase II promoter (pHIV-1/R and pHIV-1/βgl, pβgl/R and pβgl/βgl) and one pair of constructs containing an RNA polymerase III promoter (pU6/R and pU6/βgl). In cells transfected with the wild-type pHIV-1/R construct, strand-specific promoter proximal transcription was readily detected by the corresponding probe (mp18/R ss+, panel 1). In contrast, when the −5 to +82 HIV-1 sequences in this construct were replaced by the β-globin sequences (construct pHIV-1/βgl), the signal hybridizing to the corresponding probe (mp18/βgl ss+) was undetectable (panel 2). The α signals were nearly identical in both samples (probe mp18/α ss+, panels 1 and 2), indicating that the transfection efficiencies were similar. Because the levels of transcription in the absence of the −5 to +82 HIV-1 sequences were not detectable, the magnitude of the transcription stimulation by these sequences cannot be ascertained but may be very large. A similar result was obtained with the constructs pβgl/R and pβgl/βgl (panels 3 and 4). Significantly, the HIV-1 −5 to +82 region also stimulated transcription from the RNA polymerase III U6 promoter about fourfold (cf. panels 5 and 6), indicating that it contains an activator of both RNA polymerase II and RNA polymerase III transcription. As described previously by others in a similar assay (Kao et al. 1987), the presence of Tat did not increase promoter proximal transcription of the TAR-containing constructs (data not shown). These results demonstrate that the HIV-1 region from −5 to +82 contains an activator of transcription. They are consistent with the observations that substitution of these sequences results in a complete loss of the short transcripts and therefore a loss in total RNA (Fig. 6B). Together, the run-on analyses and the RNase T_1 protection analyses suggest that the HIV-1 sequences between −5 and +82 contain an inducer of short transcripts, which we refer to as IST.

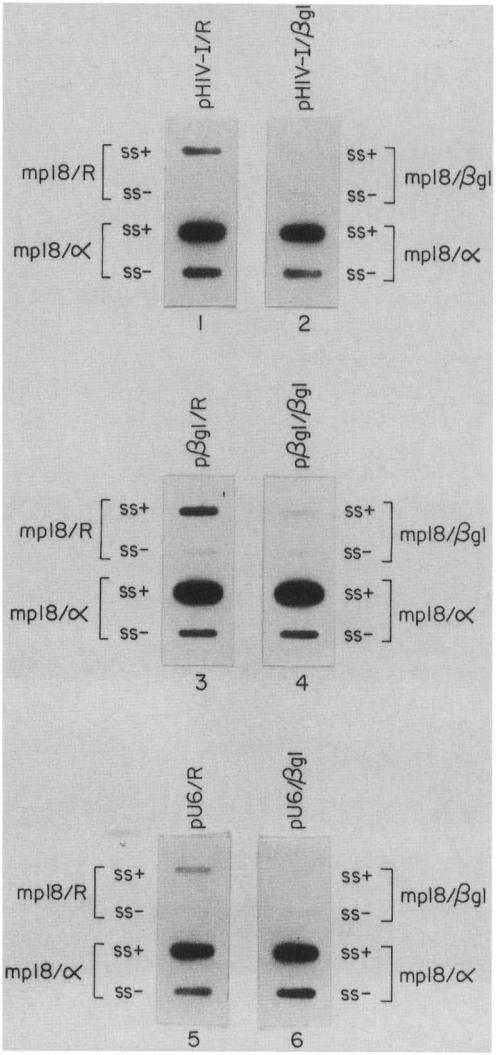


Figure 7. The HIV-1 sequences from −5 to +82 increase the amounts of newly transcribed RNA from an RNA polymerase II and an RNA polymerase III promoter. COS cells were transfected with πSVHPα2 (Treisman et al. 1983), a plasmid carrying the α-globin gene and an SV40 origin of replication, and the constructs indicated above the panels. The cells were permeabilized and incubated with [α^{32}P]UTP for 6 min. RNA was extracted and hybridized to different single-stranded probes immobilized on filters. The mp18/R probes contain the strand hybridizing to the sense (labeled ss+) or antisense (labeled ss−) strand of the HIV-1 sequences from −5 to +82, the mp18/βgl probes contain the antisense and sense strands of the rabbit β-globin gene from −5 to +82, and the mp18/α probes contain the antisense or sense strands of the whole human α-globin gene from −570 to 90 nucleotides past the polyadenylation site.

Discussion

The IST described here has a remarkably broad activity, since it is capable of inducing short transcripts from very different types of RNA polymerase II promoters (mRNA and snRNA) and even from an RNA polymerase III promoter. The activity of IST does not affect significantly the normal promoter activity that results in the synthesis of full-length transcripts, but rather superimposes itself on the normal activity of the promoter. IST increases transcription as measured by run-on experiments. This suggests that IST functions by inducing the

formation of unusual transcription complexes that are, like the U1 and U2 snRNA transcription complexes, not capable of efficient elongation.

Where is IST located?

Our experiments localize IST in the region extending from −5 to +82 relative to the HIV-1 cap site. IST could be a DNA element. Several protein-binding sites located downstream of the HIV-1 transcriptional start site have been visualized by DNase I footprinting (Garcia et al. 1987; Jones et al. 1988; Wu et al. 1988a,b), and some of these binding sites correlate with sequences required for efficient transcription in vivo and in vitro (Garcia et al. 1987; Jones et al. 1988). For example, Jones et al. (1988) identified a motif repeated three times between positions −3 and +22, which binds a factor designated LBP-1, and showed that a triple-site mutation that modifies the three LBP-1 binding sites severely reduces transcription in vitro. Primer extension, the assay used to measure transcription in these experiments, scored both short and long transcripts but could not distinguish between the two types of RNA. Thus, it is possible that the stimulation of transcription directed by the LBP-1 binding sites resulted solely in the formation of short transcripts, and that the LBP-1 binding sites represent IST. IST could also conceivably be an RNA element, and its activity might be mediated through one of the cellular factors that bind to the 5′-end region of HIV-1 RNA (Gatignol et al. 1989; Gaynor et al. 1989; Marciniak et al. 1990). IST is, however, clearly distinct from TAR, the target for Tat *trans*-activation, since a mutation (+34Δ4) that debilitates TAR does not affect the formation of short transcripts. Localizing the IST element is complicated by the fact that mutations downstream of the start site of transcription alter the sequence of the RNA and thus may change its stability. The precise localization of IST will therefore require the analysis of mutations that do not affect the secondary structure in the +2 to +59 region by both run-on and RNase T_1 protection analyses.

How does IST function?

A question related to the localization of IST is, how does IST function. IST is an activator of transcription, as demonstrated by the run-on experiments in Figure 7. However, most, if not all, of the resulting RNAs are short, as indicated by the RNase T_1 protection analyses. Thus, IST might consist of both an activator of transcription and a 3′-end formation element. However, a simpler model is that IST directs the formation of weak transcription complexes that are unable to elongate efficiently. In this model, the short RNAs would be the remnants, stabilized by the stem-and-loop structure, of exonucleolytic degradation of longer, randomly terminated RNA molecules. Random termination of transcription would occur relatively close to the start site of transcription, since IST does not increase significantly the production of polyadenylated RNA and, in our constructs, the polyadenylation site is only 236 nucleotides away from the start site of transcription. Although the secondary structure at the 5′ end of HIV-1 RNAs is not sufficient for the generation of short transcripts (Fig. 6, pHIV-1/RAT construct), it probably stabilizes the products of abortive transcription and therefore may be a key requirement for their detection.

Effects of Tat on initiation and elongation of transcription

As described previously (Rosen et al. 1985; Cullen 1986; Peterlin et al. 1986; Wright et al. 1986; Muesing et al. 1987), we find that Tat can increase the production of full-length RNAs from a wide variety of promoters, although with different efficiencies. Tat cannot, however, affect an RNA polymerase III transcription complex. In our transient transfection assay, the increase in polyadenylated RNA in the presence of Tat is accompanied by a decrease in the short transcripts. Strictly, this effect could result from Tat stabilizing the polyadenylated RNAs and destabilizing the short RNAs. However, consistent with the observations of Hauber et al. (1987), we failed to detect any effect of Tat on RNA stability as measured after blocking transcription with actinomycin D (data not shown; see Materials and methods). The increase in polyadenylated RNAs and decrease in short RNAs could also result from transcription antitermination by Tat around position +60, as proposed originally by Kao et al. (1987). However, our data do not demonstrate that the long transcripts originate from the short ones. Instead, our analysis of the pU2/R construct, which shows that in the presence of Tat new U2 transcription complexes capable of reading through the 3′ box are generated, suggests another model—that Tat directs the formation of mRNA-type transcription complexes, which, unlike the snRNA and IST-directed transcription complexes, are capable of efficient elongation through long distances. Thus, Tat action can be described as increasing the rate of initiation of mRNA-type transcription complexes, consistent with the observation that Tat affects both initiation and elongation of transcription (Laspia et al. 1989).

Possible role of the short transcripts in infected cells

An intriguing question is the role IST plays in Tat *trans*-activation. Recent results suggest that IST is not absolutely required for Tat *trans*-activation. Indeed, constructs in which most of the −5 to +82 HIV-1 sequences have been replaced by sequences encoding the RNA-binding sites for the Rev or MS2 coat proteins can be *trans*-activated by Tat-Rev or Tat-MS2 coat fusion proteins (Selby and Peterlin 1990; Southgate et al. 1990). Instead, the properties of IST suggest that this element may have a regulatory role in the virus.

The TAR element performs its function at the level of nascent RNA (Berkhout et al. 1989); it is, therefore, likely to *trans*-activate the very promoter from which it is synthesized. This implies that a given DNA template

must be transcribed to be responsive to Tat. In this context, IST provides a mechanism to allow an HIV-1 provirus to be highly responsive to Tat without itself expressing any full-length mRNAs or proteins. The Tat protein, which is a viral product, could be provided by another provirus in the same cell or perhaps even by proviruses in other cells. Indeed, Tat is secreted into the medium by HIV-1-infected cells (Ensoli et al. 1990) and might conceivably be taken up by neighboring cells. Once released from the transcription complex, short transcripts might also, under certain circumstances, down-regulate Tat *trans*-activation by sequestering Tat molecules. IST may, then, be an important player in the switch between latency and productive infection.

Because IST depends only on cellular factors for its function, it is likely to have cellular equivalents and may be just one example of a family of transcriptional elements that activate the formation of nonprocessive transcription complexes. This raises the possibility that many cellular inducible genes direct, in the absence of the inducer, the synthesis of large amounts of prematurely truncated and unstable RNA molecules and are regulated in part by transcription factors that change the elongation properties of such transcription complexes at the promoter. Thus, the HIV-1 LTR may be unique only in that the products of such abortive transcription are stable (and thus detectable), and are probably themselves involved in transcription regulation by Tat.

Materials and methods

Plasmid constructions

pHIV-1/R is a derivative of pUC118 with a large insert between the *Eco*RI and *Hin*dIII sites of the polylinker. The insert consists of (1) a duplicated SV40 early polyadenylation site at which RNAs derived from cryptic promoters within the vectors are presumably cleaved and polyadenylated (Hernandez and Lucito 1988), (2) HIV-1 sequences from −138 to +82 relative to the cap site but modified from −10 to −6 by a clustered point mutation that creates a *Xho*I restriction site from −10 to −5, (3) a fragment derived from Ad2 (nucleotides 22,305 to 22,491) carrying the L3 polyadenylation site, (4) a 214-nucleotide *Hinc*II–*Pvu*II fragment derived from pBSM13(+) carrying the T3 promoter, and (5) an SV40 fragment (nucleotides 5171 to 105) carrying the origin of replication (but not the SV40 enhancer). The promoter fragments replacing the HIV-1 promoter sequences from an *Eco*RI site (see Fig. 1A) at position −140 to the *Xho*I site were as follows: for pβgl/R, a *Xho*II–*Pvu*II rabbit β-globin promoter fragment extending from −426 to −9 relative to the cap site (the pβgl/R clone used in Fig. 6 has one additional nucleotide at the junction between β-globin and HIV-1 sequences GGGCAG**C**TCGAGTACTG as compared to the clone used in Fig. 2, GGGCAGTCGAGTACTG, and this results in the generation of two transcriptional start sites); for pMLP/R, an Ad2 fragment extending from −188 to −12 relative to the major late cap site; for pαgl/R, a *Bss*HII human α-globin promoter fragment extending from −486 to −14 relative to the cap site; for pU2/R, a human U2 promoter fragment extending from −247 to −5 relative to the cap site; for pU6/R, a human U6 promoter fragment extending from −245 to −11. In pHIV-1/R/2 and pU2/R/2, a fragment derived from pU6/Hae/RA.2 (Lobo and Hernandez 1989) containing the last 6 nucleotides of the coding U1 region and 90 nucleotides of U1 3′-flanking sequences was inserted between the HIV-1 sequences and the Ad2 fragment, at the *Bam*HI site. The +34Δ4 mutation was introduced into several of the constructs by cleavage with *Sac*I, digestion of the 3′ overhangs with Klenow fragment, and religation. The clustered point mutations in the proximal elements of the U2 and U6 promoters were introduced by oligonucleotide-directed mutagenesis and were equivalent to the LS−34/−43 (Hernandez and Lucito 1988) and LS2 (Lobo and Hernandez 1989) mutations described previously.

To generate pHIV-1/βgl, rabbit β-globin sequences from −4 to +82 were amplified from a plasmid by the polymerase chain reaction (PCR) with two primers; the upstream primer contained an *Xho*I site, the downstream primer contained a *Bam*HI site, such that after digestion with *Xho*I and *Bam*HI, the fragment could be inserted by "sticky-end" ligation into the pHIV-1/R construct to replace the HIV-1 −5 to +82 sequences (see Fig. 1A for the location of the *Xho*I and *Bam*HI restriction sites). To construct pβgl/βgl, rabbit β-globin sequences from −353 to +82 were amplified and cloned by similar methods as described above between the *Eco*RI and *Bam*HI sites. pHIV-1/RAT was constructed from two oligonucleotides; the upper-strand oligonucleotide had the sequence GGAACTCGAGTACTGCCAAGGGATCGATCGGTCTCTCGAGGGTCCGAGTCTAG, and the lower-strand oligonucleotide had the sequence AAAAGGATCCCCAGCTTTATTGAGGCTTAAGCAGTGCCAGAGAGACCAATCTGGTCTAGACTCGGACCCTCGA. The two oligonucleotides were annealed, the ends were filled in with Klenow, and the resulting fragment was digested with *Bam*HI and partially digested with *Xho*I, and then cloned by sticky-end ligation into HIV-1/R to replace the −5 to +82 HIV-1 sequences.

The pCG expression vector has been described previously (Tanaka and Herr 1990). To construct pCGTat, Tat cDNA sequences from the initiation to the termination codon were amplified by PCR with an upstream primer containing an *Xba*I site and a downstream primer containing a *Bam*HI site. The resulting fragment was cleaved with *Xba*I and *Bam*HI and cloned by sticky-end ligation into pCG. All clones were verified by sequencing.

Transfections

Jurkat, HeLa, and COS cells were transfected by electroporation. Jurkat cells were grown in RPMI-1640 with 10% fetal calf serum, 10 mM HEPES (pH 7.4), and 2 mM L-glutamine, to a density of 1×10^6 to 1.5×10^6 cells per milliliter. For each sample, cells from 15 ml of culture were collected by centrifugation and resuspended in 250 μl of medium. A mixture of plasmids containing 7 μg of the test plasmid, 1 or 2 μg of pα1x72, 0.5 μg of pCG or pCGTat, and calf thymus DNA carrier to a total of 20 μg in 20–30 μl of TE (10 mM Tris/HCl at pH 7.5, 1 mM EDTA) was added to the cells. The mixture was incubated on ice for 10 min, transferred to a Gene Pulser Cuvette (Bio-Rad, 0.4-cm electrode gap), and electroporated by a pulse of 200 V, 960 μF in a Gene Pulser (Bio-Rad). The cells were then transferred to 10-cm dishes containing 10 ml of medium, and RNA was collected 24 hr later. HeLa cells and COS cells were grown in DMEM supplemented with 10% fetal calf serum to a density of 80–90%. For each sample, the cells from one 10-cm dish were trypsinized, resuspended in 250 μl of medium, and electroporated as described above. Where indicated in the figure legends, 3 μg of pSVEori−, an SV40 T Ag-expressing vector (Hernandez and Lucito 1988), was included in the transfections.

Determination of RNA stabilities

To determine the stabilities of the different RNA species, Jurkat cells were transfected as described above. Twenty-two

hours after transfection, actinomycin D was added to the medium to a final concentration of 2.5 μg/ml to block all transcription. An aliquot of cells was collected immediately, and further aliquots were collected after different periods of time. RNA was analyzed by RNase T_1 as described below. The half-lives determined in such experiments were >7 hr for both the short transcripts and polyadenylated RNAs. In contrast, RNAs ending at the U1 3′ box or at the run of T residues were quite unstable, with a half-life of 1 hr. The half-lives were not changed in the absence or presence of Tat.

RNase T_1 mapping and primer extension

RNA was collected as described in Hernandez and Lucito (1988). The RNase T_1 protections and the primer extensions were performed as described in Lobo and Hernandez (1989). To generate the riboprobes, each different construct was cleaved with *Eco*RI and used as a template for bacteriophage T3 RNA polymerase. Gels were quantitated by liquid scintillation counting of excised bands and normalization to the α-globin signal from the internal reference.

Run-on analyses

Five 10-cm dishes of 80–90% confluent COS cells were transfected separately by electroporation as described above. The cells were then mixed and transferred to three 15-cm dishes. Twenty-four hours after transfection, the cells were permeabilized by treatment with digitonin according to Ucker and Yamamoto (1984). Permeabilized cells were resuspended in 100 μl of buffer B (50 mM HEPES at pH 8, 5 mM $MgCl_2$, 0.5 mM DTT, 25% glycerol). The reaction was started by the addition of 35 μl of [α-^{32}P]UTP (350 μCi) at a sp. act. of 3000 Ci/mmole, and 200 μl of a solution containing 0.3 M NH_4Cl, 0.5 mM HEPES (pH 8), 6 mM MgO acetate, 4 mM $MgCl_2$, 4 mM DTT, 0.8 mM ATP, CTP, GTP, and 0.5 μM UTP, and the mixture was incubated for 6 min at 30°C. The reaction was terminated by the addition of 600 μl of a solution containing 0.5 M NaCl, 50 mM $MgCl_2$, 2 mM $CaCl_2$, 10 mM Tris/HCl (pH 7.4), 5 mM DTT, 120 units/ml of DNase I, and 2 mg/ml of α-amanitin. After 10 min at 30°C, 200 μl of SDS/Tris buffer was added (5% SDS, 0.5 M Tris/HCl at pH 7.4, 0.125 M EDTA), and the RNA was extracted as described by Greenberg and Bender (1987).

The radiolabeled RNA (7×10^6 to 10×10^6 cpm) was hybridized to single-stranded DNA probes immobilized on filters (Biodyne transfer membrane) in 1× Denhardt's solution containing 100 μg/ml of denatured salmon sperm DNA, 0.5% SDS, and 6× SSC for 48 hr at 55°C. The filters were washed, treated with RNase A (10 μg/ml in 6× SSC) for 30 min at 37°C, rinsed in 6× SSC at 37°C, blotted dry, and autoradiographed. The results were quantitated by liquid scintillation counting of the slots and normalization to the α-globin RNA signal from the internal reference.

Acknowledgments

We thank J.D. Watson for encouraging us to start this work, W. Phares for teaching us how to transfect Jurkat cells, M. Laspia, and M. Kessler for help with the run-on experiments, S. Ifill for help with the cloning, M. Gilman, W. Herr, M. Laspia, S. Lobo, W. Phares, J. Skowronski, and M. Tanaka for stimulating discussions and gifts of plasmids, and W. Herr, M. Laspia, S. Lobo, M. Mathews, and B. Stillman for comments on the manuscript. We also thank M. Ockler, J. Duffy, and P. Renna for artwork and photography. This work was funded by grants 5PO1 AI27270-02 and ROI GM38810-03 from the National Institutes of Health (NIH). M.S. was supported initially by the Genetics Training Grant GM07964 from the NIH. N.H. is a Rita Allen Foundation Scholar.

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

References

Bentley, D.L. and M. Groudine. 1988. Sequence requirements for premature termination of transcription in the human c-myc gene. *Cell* **53:** 245–256.

Berkhout, B., R.H. Silverman, and K.-T. Jeang. 1989. Tat *trans*-activates the human immunodeficiency virus through a nascent RNA target. *Cell* **59:** 273–282.

Cullen, B.R. 1986. *Trans*-activation of human immunodeficiency virus occurs via a bimodal mechanism. *Cell* **46:** 973–982.

Dahlberg, J.E. and E. Lund. 1988. The genes and transcription of the major small nuclear RNAs. In *Structure and function of major and minor small nuclear ribonucleoprotein particles* (ed. M. Birnstiel), pp. 38–70. Springer-Verlag, Heidelberg.

Das, G., D. Henning, D. Wright, and R. Reddy. 1988. Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III. *EMBO J.* **7:** 503–512.

Ensoli, B., G. Barillari, S.Z. Salahuddin, R.C. Gallo, and F. Wong-Staal. 1990. Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature* **345:** 84–86.

Feng, S. and E.C. Holland. 1988. HIV-1 *tat trans*-activation requires the loop sequence within *tar*. *Nature* **334:** 165–167.

Garcia, J.A., F.K. Wu, R. Mitsuyasu, and R.B. Gaynor. 1987. Interactions of cellular proteins involved in the transcriptional regulation of the human immunodeficiency virus. *EMBO J.* **6:** 3761–3770.

Gatignol, A., A. Kumar, A. Rabson, and K.-T. Jeang. 1989. Identification of cellular proteins that bind the human immunodeficiency virus type 1 trans-activation-responsive TAR element RNA. *Proc. Natl. Acad. Sci.* **86:** 7828–7832.

Gaynor, R., E. Soultanakis, M. Kuwabara, J. Garcia, and D.S. Sigman. 1989. Specific binding of a HeLa nuclear protein to RNA sequences in the human immunodeficiency virus transactivating region. *Proc. Natl. Acad. Sci.* **86:** 4858–4862.

Greenberg, M.E. and T.P. Bender. 1987. Identification of newly transcribed RNA. In *Current protocols in molecular biology* (ed. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl), p. 4.10.3, Greene Publishing Associates/Wiley-Interscience, New York.

Hauber, J., A. Perkin, E.P. Heimer, and B.R. Cullen. 1987. Trans-activation of human immunodeficiency virus gene expression is mediated by nuclear events. *Proc. Natl. Acad. Sci.* **84:** 6364–6368.

Hauber, J. and B.R. Cullen. 1988. Mutational analysis of the *trans*-activation-responsive region of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* **62:** 673–679.

Hernandez, N. 1985. Formation of the 3′ end of U1 snRNA is directed by a conserved sequence located downstream of the coding region. *EMBO J.* **4:** 1827–1837.

Hernandez, N. and R. Lucito. 1988. Elements required for transcription initiation of the human U2 snRNA gene coincide with elements required for snRNA 3′ end formation. *EMBO J.* **7:** 3125–3134.

Hernandez, N. and A.M. Weiner. 1986. Formation of the 3′ end

of U1 snRNA requires compatible snRNA promoter elements. *Cell* **47:** 249–258.

Jakobovits, A., D.H. Smith, E.B. Jakobovits, and D.J. Capon. 1988. A discrete element 3′ of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by HIV *trans*-activator. *Mol. Cell. Biol.* **8:** 2555–2561.

Jones, K.A., P.A. Luciw, and N. Duchange. 1988. Structural arrangements of transcription control domains within the 5′-untranslated leader regions of the HIV-1 and HIV-2 promoters. *Genes Dev.* **2:** 1101–1114.

Kao, S.-Y., A.F. Calman, P.A. Luciw, and B.M. Peterlin. 1987. Antitermination of transcription within the long terminal repeat of HIV by the *tat* gene product. *Nature* **330:** 489–493.

Laspia. M.F., A.P. Rice, and M.B. Mathews. 1989. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. *Cell* **59:** 283–292.

Lobo, S.M. and N. Hernandez. 1989. A 7 bp mutation converts a human RNA polymerase II snRNA promoter into an RNA polymerase III promoter. *Cell* **58:** 55–67.

Marciniak, R.A., M.A. Garcia-Blanco, and P.A. Sharp. 1990. Identification and characterization of a HeLa nuclear protein that specifically binds to the trans-activation response (TAR) element of human immunodeficiency virus. *Proc. Natl. Acad. Sci.* **87:** 3624–3628.

Mattaj, I.W., N.A. Dathan, H.W. Parry, P. Carbon, and A. Krol. 1988. Changing the RNA polymerase specificity of U snRNA gene promoters. *Cell* **55:** 435–442.

Muesing, M.A., D.H. Smith, and D.J. Capon. 1987. Regulation of mRNA accumulation by a human immunodeficiency virus *trans*-activator protein. *Cell* **48:** 691–701.

Neuman de Vegvar, H.E., E. Lund, and J.E. Dahlberg. 1986. 3′ End formation of U1 snRNA precursors is coupled to transcription from snRNA promoters. *Cell* **47:** 259–266.

Neuman de Vegvar, H.E. and J.E. Dahlberg. 1989. Initiation and termination of human U1 RNA transcription requires the concerted action of multiple flanking elements. *Nucleic Acids Res.* **17:** 9305–9318.

Peterlin, B.M., P.A. Luciw, P.J. Barr, and M.D. Walker. 1986. Elevated levels of mRNA can account for the trans-activation of human immunodeficiency virus. *Proc. Natl. Acad. Sci.* **83:** 9734–9738.

Rice, A.P. and M.B. Mathews. 1988. Transcriptional but not translational regulation of HIV-1 by the *tat* gene product. *Nature* **332:** 551–553.

Rosen, C.A., J.G. Sodroski, and W.A. Haseltine. 1985. The location of *cis*-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. *Cell* **41:** 813–823.

Rougvie, A.E. and J.T. Lis. 1988. The RNA polymerase II molecule at the 5′ end of the uninduced *hsp70* gene of D. melanogaster is transcriptionally engaged. *Cell* **54:** 795–804.

Roy, S., N.T. Parkin, C. Rosen, J. Itovitch, and N. Sonenberg. 1990. Structural requirements for *trans*-activation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by *tat*: Importance of base pairing, loop sequence, and bulges in the *tat*-responsive sequence. *J. Virol.* **64:** 1402–1406.

Sadaie, M.R., T. Benter, and F. Wong-Staal. 1988. Site-directed mutagenesis of two trans-regulatory genes (*tat*III, *trs*) of HIV-1. *Science* **239:** 910–913.

Selby, M.J. and B.M. Peterlin. 1990. Transactivation by HIV-1 Tat via a heterologous RNA-binding protein. *Cell* **62:** 769–776.

Selby, M.J., E.S. Bain, P.A. Luciw, and B.M. Peterlin. 1989. Structure, sequence, and position of the stem-loop in *tar* determine transcriptional elongation by *tat* through the HIV-1 long terminal repeat. *Genes Dev.* **3:** 547–558.

Southgate, C., M.L. Zapp, and M.R. Green. 1990. Activation of transcription by HIV-1 Tat protein tethered to nascent RNA through another protein. *Nature* **345:** 640–642.

Spencer, C.A. and M. Groudine. 1990. Transcription elongation and eucaryotic gene regulation. *Oncogene* **5:** 777–785.

Spencer, C.A., R.C. LeStrange, U. Novak, W.S. Hayward, and M. Groudine. 1990. The block to transcription elongation is promoter dependent in normal and Burkitt's lymphoma c-myc alleles. *Genes Dev.* **4:** 75–88.

Steitz, J.A., D.L. Black, V. Gerke, K.A. Parker, A. Kramer, D. Frendewey, and W. Keller. 1988. Functions of the abundant U-snRNPs. In *Structure and function of major and minor small nuclear ribonucleoprotein particles* (ed. M. Birnstiel), pp. 115–154. Springer-Verlag, Heidelberg.

Tanaka, M. and W. Herr. 1990. Differential transcriptional activation by Oct-1 and Oct-2: Interdependent activation domains induce Oct-2 phosphorylation. *Cell* **60:** 375–386.

Treisman, R., M.R. Green, and T. Maniatis. 1983. Cis and trans activation of globin gene transcription in transient assays. *Proc. Natl. Acad. Sci.* **80:** 7428–7432.

Ucker, D.S. and K.R. Yamamoto. 1984. Early events in the stimulation of mammary tumor virus RNA synthesis by glucocorticoids. *J. Biol. Chem.* **259:** 7416–7420.

Wright, C.M., B.K. Felber, H. Paskalis, and G.N. Pavlakis. 1986. Expression and characterization of the trans-activator of HTLV-III/LAV virus. *Science* **234:** 988–992.

Wu, F., J. Garcia, R. Mitsuyasu, and R. Gaynor. 1988a. Alterations in binding characteristics of the human immunodeficiency virus enhancer factor. *J. Virol.* **62:** 218–225.

Wu, F., J.A. Garcia, D. Harrich, and R.B. Gaynor. 1988b. Purification of the human immunodeficiency virus type 1 enhancer and TAR binding proteins EBP-1 and UBP-1. *EMBO J.* **7:** 2117–2129.

Zucker, M. and P. Stiegler. 1981. Optimal computer folding of large RNA sequences using thermodynamic and auxiliary information. *Nucleic Acids Res.* **9:** 133–148.