JOURNAL OF VIROLOGY, July 1991, p. 3468–3474
0022-538X/91/073468-07$02.00/0

Vol. 65, No. 7

Mutational Analysis of the Equine Infectious Anemia Virus Tat-Responsive Element

MAGDA CARVALHO AND DAVID DERSE*

Laboratory of Viral Carcinogenesis, National Cancer Institute, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702-1201

Received 28 January 1991/Accepted 25 March 1991

A hairpinlike structure is predicted to exist at the 5′ end of equine infectious anemia virus (EIAV) RNA which is similar in many ways to the human immunodeficiency type 1 (HIV-1) Tat-responsive element (TAR). In EIAV, this structure has a shorter stem than in HIV-1 and lacks the uridine bulge. Primer extension analysis of EIAV RNA was used to identify the transcriptional start site in the viral long terminal repeat. Premature termination of primer elongation at the predicted double-stranded RNA region was frequently observed and suggests that the inferred hairpin structure exists under these conditions. We have functionally characterized EIAV TAR by site-directed mutagenesis and transient gene expression analysis. It is demonstrated here that the secondary structure of this element is essential for Tat action. Mutations that disrupted base pairing abolished TAR function, and compensatory mutations that restored the stem structure resulted in Tat activation. The TAR loop appears to be closed by two U · G base pairs that are likely to provide a unique structural motif recognized by the Tat protein. With one exception, substitutions of nucleotides within the EIAV loop sequence decreased TAR function. All nucleotide substitutions of the cytidine at position +14 increased EIAV Tat responsiveness; however, its deletion abolished *trans* activation. Our results lead us to propose that the EIAV and HIV-1 Tat systems employ closely related *cis*- and *trans*-acting components that probably act by the same mechanism.

Equine infectious anemia virus (EIAV) is the etiologic agent of equine infectious anemia, a relapsing disease of horses characterized by recurrent episodes of clinical illness and progressive development of immune-mediated lesions (5). EIAV belongs to the lentivirinae subfamily of retroviruses. The lentiviruses share similarities in genome organization, biologic properties, and the propensity for persistent infection in the natural host. The other members of this group include the human (HIV) and simian immunodeficiency viruses, visna virus, caprine arthritis-encephalitis virus, feline and bovine immunodeficiency-like viruses. Like the other lentiviruses, EIAV infects cells of the monocyte macrophage lineage (19).

One striking feature of the lentivirus genome is the presence of several accessory genes in addition to *gag*, *pol*, and *env* (7, 27). HIV has at least six additional genes located either between *pol* and *env* or at the 3′ end of the genome, while the less complex genome of EIAV has three additional open reading frames arranged in a similar manner (10, 31, 36). The replication of HIV-1 has been shown to be dependent on the expression of the *trans*-activator protein encoded by the *tat* gene of the virus (1, 35). EIAV and primate lentivirus Tat proteins show two regions of highly conserved amino acid sequence (10, 36); these include a core region thought to be an essential part of the activation domain (4, 26, 33) and a basic region believed to be necessary for Tat-responsive element (TAR) RNA binding (4, 12, 34, 37). The structural similarities between EIAV and HIV Tat proteins suggest that they act by similar mechanisms. It should be noted, however, that EIAV and HIV-1 Tat proteins will not activate the heterologous long terminal repeat (LTR) (10).

HIV-1 Tat increases the steady-state levels of viral transcripts by interacting with the *cis*-acting TAR present at the 5′ end of viral RNA (6, 17, 23, 28, 35). It is likely that binding to TAR RNA positions Tat for subsequent interactions with the transcription apparatus. HIV-1 TAR assumes a stable stem-loop structure; mutagenesis studies have defined functionally important regions of this element for both *trans* activation in vivo and Tat binding in vitro (2, 3, 9, 13, 14, 17, 18, 32). It has been shown that HIV-1 Tat binds specifically to a part of the double-stranded stem that contains a bulge of unpaired uridines (8, 29, 39). However, mutations in the loop that had no effect on Tat binding in vitro (29, 30, 39) reduced *trans* activation in vivo (3, 13, 14). Thus, cellular factors may be involved in Tat-mediated *trans* activation, and nuclear factors that bind specifically to TAR RNA have been identified (15, 16, 21). The roles of cellular TAR-binding proteins remain to be defined.

It has been shown that sequences downstream of the RNA start site are necessary for EIAV transactivation (11) and that the 5′ end of EIAV RNA can be folded to form a stem-loop structure similar in some respects to the TAR of primate lentiviruses. The structural similarities of EIAV and HIV Tat proteins combined with the observation that their *cis*-acting elements also share relevant features provide the basis for our proposal that the mechanism of Tat action is similar in EIAV and HIV-1. We have characterized EIAV TAR by mutagenesis of individual bases and by substitution of clusters of nucleotides within the LTR. The mutated LTRs were coupled to the bacterial chloramphenicol acetyltransferase (*cat*) gene, and their transcriptional activities were tested in transiently transfected cells. These studies indicated that EIAV TAR is a structured RNA element; the secondary structure of the stem, unique U · G base pairs adjacent to the loop, and specific nucleotides within the loop are critical determinants of Tat-mediated *trans* activation.

* Corresponding author.

We conclude that despite their great divergence, HIV and EIAV appear to have similar mechanisms of gene regulation.

MATERIALS AND METHODS

Plasmids. Plasmids pRS-Etat-M, pRSPA, pUX-E478 (parental EIAV LTR), and pUXCAT used in these experiments have been previously described (10, 11). pRSPA is an expression vector with a Rous sarcoma virus promoter and simian virus 40 polyadenylation signal in a Bluescript KS+ (Stratagene) backbone. pRS-Etat-M contains the EIAV *tat* gene with a synthetic methionine initiation codon cloned in pRSPA. pUXCAT contains the bacterial *cat* gene with simian virus 40 polyadenylation signals in a modified pUC18 plasmid; pUX-E478 contains the EIAV LTR cloned in pUXCAT.

Mutagenesis. The EIAV LTR from pUX-E478 was cloned into the Bluescript KS+ polylinker to give pKS-E-LTR, which was used as the mutagenesis substrate. Our mutants were obtained by the following approaches.

Oligonucleotide-directed mutagenesis. pE-12C-A and pE-17G-U were obtained by oligonucleotide-directed mutagenesis. pKS-E-LTR was used to transform the bacterial strain CJ236 (Bio-Rad), and single-stranded DNA was produced by superinfection with the helper phage VCSM13 (Stratagene). Single-stranded DNA was isolated and purified, and second-strand synthesis was performed essentially as described previously (20), using the oligonucleotides 5′-CAGATTA<u>A</u>TGCGGTCTGA-3′ and 5′-CAGATTCTGCG<u>T</u>TCTGA-3′, containing the underlined single point mutations as primers. The double-stranded reaction products were isolated and used to transform the bacterial strain JM101. Mutants were characterized by dideoxynucleotide sequencing. The mutated plasmids pKS-12C-A and pKS-17G-U were digested with *Hin*dIII; the LTR-containing fragments were purified by polyacrylamide gel electrophoresis and ligated into the *Hin*dIII restriction site of pUXCAT.

Primer-directed mutagenesis by *in vitro* amplification of EIAV LTR fragments. pE-14G-C, pE-16G-A, and pE-10,11UU were obtained by in vitro amplification of the EIAV LTR R region, using a mutagenic oligonucleotide primer and ligating this to a U3 region fragment. pKS-E-LTR was digested with *Pvu*II prior to amplification. The R regions were amplified by using the 5′ oligonucleotide primers (Operon) 5′-GGGCACTCAGATTCT<u>C</u>CGGTCTGAGTC-3′, 5′-GGGCACTCAGATTCTGC<u>A</u>GTCTGAGTC-3′, and 5′-GGGCACTCAGACTGCGGTCTGAGTC-3′ to introduce the desired substitutions or, in the last case, a deletion. The KS sequencing primer (Stratagene) served as the 3′ primer. Amplification reactions (50 μl) were carried out with 200 ng of DNA template, 200 ng of each primer, 200 μM deoxynucleoside triphosphate, each 1.5 U of *Taq* polymerase (Perkin-Elmer), and buffer (50 mM KCl, 20 mM Tris-HCl [pH 8.3], 2 mM $MgCl_2$, 0.1 mg of bovine serum albumin per ml). The reactions were subjected to 94°C for 5 min for DNA denaturation, followed by 30 cycles of 94°C for 30 s, 55°C for 1 min, and 72°C for 2 min and finishing with 72°C for 10 min. The amplified products were purified by polyacrylamide gel electrophoresis and cloned into pUXCAT (previously cut with *Sma*I and *Hin*dIII). The U3 region was amplified by using the reaction conditions described above. The reverse sequencing primer served as the 5′ primer, and the 3′ primer was the U3 region-specific oligonucleotide 3′-GTGTTG TATTCTGACAATT-5′. The U3 region amplification product was digested with *Hin*dIII and ligated to the R-region plasmids that had been cut with *Sma*I, thus fusing the blunt ends of the R and U3 regions. As these LTRs were cloned into the *Sma*I site of pUXCAT, they have at their 3′ ends 32 nucleotides from the polylinker not present in other constructs.

The remainder of the constructs were obtained by joining independently amplified U3 and R regions by overlap amplification prior to cloning into pUXCAT. pKS-E-LTR was digested with *Pvu*II prior to amplification. pE-Stem(+19/+24) was made by first amplifying the U3 region, using the reverse-sequencing primer and an EIAV LTR primer, 3′-CCCGTGAGTCTAAGACGCCA-5′. The R region was amplified by using a primer pair consisting of the vector-complementary sequencing primer and an EIAV primer containing a six-nucleotide alteration, 5′-ATTCTGCGGT <u>TCAGAC</u>CCCTTCTCTGCTGG-3′. The two EIAV primers possess 11 complementary bases at their 5′ ends. The products of the two amplification reactions were gel purified, mixed, and amplified by using the vector-targeted sequencing and reverse-sequencing primer pair. The polymerase chain reaction mixture was subjected to 1 cycle of 94°C for 5 min, 30°C for 1 min, and 55°C for 30 min; 30 cycles of 94°C for 0.5 min, 50°C for 1 min, and 72°C for 2 min; and 1 cycle of 72°C for 10 min. Thus, the entire LTR with six-nucleotide internal mutations was regenerated by overlap amplification. This product was digested with *Hin*dIII, gel purified, and ligated to pUXCAT. pE-Stem(+3/+8−+19/+24) was made in a similar manner; the upstream region was amplified by using reverse sequencing primer and an oligonucleotide, 3′-AGACTGTTAACCCG<u>CAGACT</u>TAAGACGCCA-5′, that introduces a six-base alteration in the LTR. The product of this reaction was gel purified and combined with the product of the R-region amplification used to generate pE-Stem(+19/+24); these plasmids were amplified by using the sequencing and reverse-sequencing primers and cloned as before.

The remaining plasmids with single mutations in the proposed loop region, pE-10U-G, pE-11U-G, pE-10,11GG, pE-17-A-18, pE-13U-A, pE-15C-U, pE-15C-G, pE-15C-A, and pE-15C-O, were also made by overlap amplification. For all of these plasmids, the upstream region was amplified by using the reverse-sequencing primer and the oligonucleotide 3′-CTGTTAACCCGTGAGTCT-5′. The R regions were amplified by using the sequencing primer and an EIAV primer such as 5′-GGGCACTCAGATTCTG<u>G</u>GGTCTGAGTC-3′ that contains the single base alteration and 11 bases complementary to the upstream fragment. The two amplified products were combined and amplified by using the two vector-targeted primers. The LTR-containing fragments were gel purified, cut with *Hin*dIII, and ligated into pUXCAT polylinker. All plasmids were banded twice in CsCl density gradients. Two independent clones of each plasmid were used in transfections. All plasmids were confirmed by double-stranded dideoxynucleotide sequencing.

Transfections and CAT assays. D17 (canine osteosarcoma) cells were maintained in Dulbecco's modified Eagle medium containing 10% fetal calf serum. The day before transfection, cells were seeded at 3×10^5 per well in 3-cm-diameter, six-well dishes. Cells were transfected with 1 μg of expression plasmid and 3 μg of CAT reporter construct by the calcium phosphate method. Four hours after transfection, cells were washed in phosphate-buffered saline and refed with fresh medium. Cell monolayers were harvested for CAT assays by scraping at 44 h after transfection. Cell extracts were prepared as described previously (10, 11), and CAT assays were performed by the solvent partition method (24).

RNA extraction and primer extension analysis. Cells were

plated and transfected as described above, collected 44 h after transfection by scraping, and transferred to 1.5-ml centrifuge tubes. After centrifugation for 15 s, the supernatant was removed and cells were disrupted in 100 μl of cold lysis buffer (10 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1.5 mM $MgCl_2$, Nonidet 0.5% P-40). Tubes were maintained on ice for 10 min and then centrifuged to pellet nuclei. To the supernatant was added 100 μl of 100 mM sodium acetate–2% sodium dodecyl sulfate and then 200 μl of phenol (equilibrated with 50 mM sodium acetate [pH 5.2]). After vigorous vortexing, the solution was incubated at 65°C for 15 min and then put on ice for 15 min. RNA in the aqueous phase was ethanol precipitated and collected by centrifugation, and the dried pellet was dissolved in 30 μl of water that had been treated with diethyl pyrocarbonate. Ten microliters of this RNA solution was used in the primer extension reactions. The oligonucleotide primer, complementary to sequences at the 5′ end of the *cat* gene, was end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. Primer extension reactions were performed as previously described (38). Following the reaction, 100 μl of 10 mM Tris-HCl–0.1 M NaCl–0.5 mM EDTA was added, and the reaction mixture was extracted with 100 μl of phenol-chloroform-isoamyl alcohol (20:20:1). The aqueous phase was ethanol precipitated, and the dried pellet was dissolved in loading buffer. After denaturation at 90°C for 5 min, the products were separated by electrophoresis on a 6% polyacrylamide–7 M urea sequencing gel.

RESULTS

Proposed secondary structure of EIAV TAR. Our initial working hypothesis was that the EIAV Tat protein mediates *trans* activation via a *cis*-acting TAR in the form of an RNA stem-loop structure. It was previously shown that sequences downstream of the RNA start site are required for EIAV Tat activation (11), and computer-assisted folding revealed that nucleotides +1 to +25 of EIAV RNA can form a stem-loop structure (Fig. 1). This structure is similar to but distinct from the structures present in HIV-1 and other primate lentiviruses (Fig. 1). EIAV TAR consists of nucleotides +1 to +25 and is considerably smaller than that of HIV-1, which contains bases spanning a region from +1 to +60. However, it is interesting to note that the minimal biologically active part of HIV-1 TAR is contained between positions +18 and +44 (14), similar in size to EIAV TAR. Although the hairpinlike structures of EIAV and HIV-1 TARs are similar, the unpaired pyrimidine bulge in the stem of HIV-1 TAR is absent from the stem of EIAV TAR. The two uridines present in EIAV TAR at positions +10 and +11 probably base pair with the two guanosines at positions +16 and +17 (25) to close the four-base loop.

To test our hypothesis directly, we constructed a series of EIAV LTR mutations that specifically disrupted the proposed stem structure or altered bases within the loop. Most of the mutations were introduced into the EIAV LTR by primer-directed mutagenesis using the polymerase chain reaction. These LTRs were linked to the *cat* gene and were tested for promoter activity in transient expression assays in D17 cells transfected with or without the EIAV *tat* expression plasmid (pRS-Etat-M). At 44 h after transfection, the cells were harvested and cell lysates were assayed for either CAT enzyme activity or CAT mRNA.

CAT activity directed by the mutated EIAV LTRs. Figure 2 presents the nucleotide alterations introduced into the putative EIAV TAR and summarizes their relative *trans*-activation efficiencies obtained from CAT assays. Levels of CAT enzyme activity are graphically summarized in Fig. 3. In the absence of EIAV Tat, the wild-type and mutated LTRs directed comparable levels of CAT activity (data not shown). The wild-type EIAV LTR was activated approximately 40 times by the addition of EIAV Tat (Fig. 3).

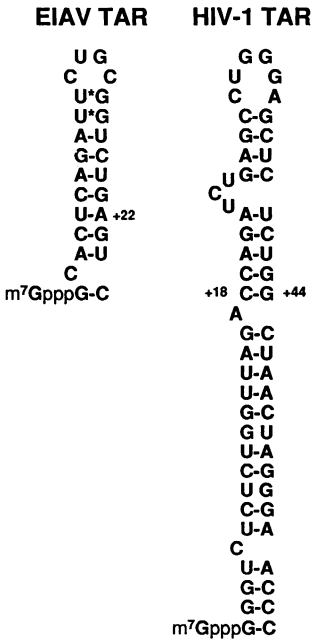

FIG. 1. Predicted secondary structure for EIAV TAR RNA and reported secondary structure of HIV-1 TAR RNA. Numbers refer to nucleotide positions with respect to the start of transcription (+1). The non-Watson-Crick U · G base pairs in EIAV TAR are indicated with asterisks.

To examine the biologic significance of the putative secondary structure, we introduced a clustered mutation that disrupts potential base pairing. The plasmid, pE-Stem(+19/+24), had a six-nucleotide alteration within the predicted stem between positions +19 and +24 (Fig. 2) and failed to respond to EIAV Tat *trans* activation. Base pairing was restored in pE-Stem(+3/+8−+19/+24) by introducing a complementary sequence which replaced the six nucleotides between positions +3 and +8 (Fig. 2A). This compensatory mutation restored *trans* activation; interestingly, CAT activity was higher than that of the wild-type LTR (Fig. 3). These results show that the presence of a secondary structure, but not the primary sequence in this region, is essential for EIAV Tat-mediated *trans* activation.

To assess the importance of the two U · G base pairs at the base of the EIAV TAR loop, we made the alterations presented in Fig. 2B. Tat responsiveness of pE-10,11UU, in which the two uridines were deleted, was almost completely abolished (Fig. 3). Likewise, substitution of the two uridines with guanines in pE-10,11-GG, such that U · G base pairs were abolished but the spacing was preserved, severely reduced TAR function. This result demonstrated that the two uridines are essential for Tat-mediated *trans* activation. Substitution of the uridine at position +11 alone, as in pE-11U-G, actually increased TAR function compared with the wild type. Thus, a single U · G base pair at the base of the loop appears to be sufficient for TAR recognition. Construct pE-17-A-18 has an additional adenosine inserted between positions +17 and +18 and may form a base pair

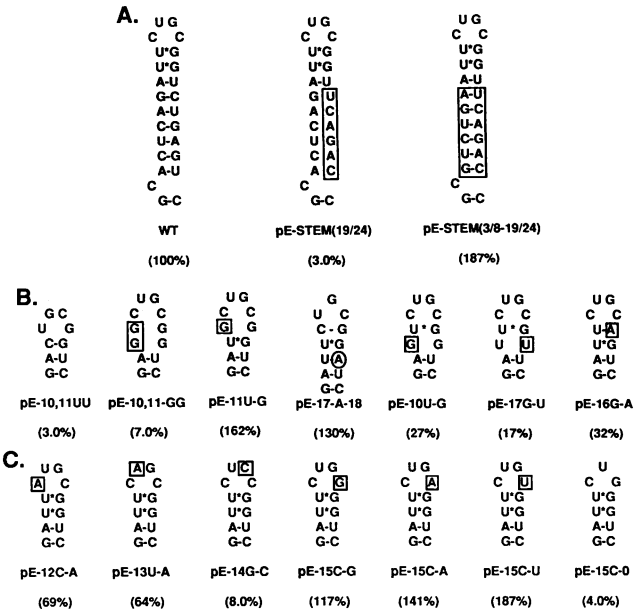

FIG. 2. Summary of stem (A), U · G base pairs (B), and loop (C) mutations in EIAV TAR. The numbers given to the constructs indicate the position of the altered nucleotide relative to the cap site (+1) followed by the base substitution. Nucleotide changes are boxed, and the single base insertion is circled. The levels of Tat-mediated *trans* activation, as percentages of wild-type LTR activity determined by CAT assay, are shown in parentheses.

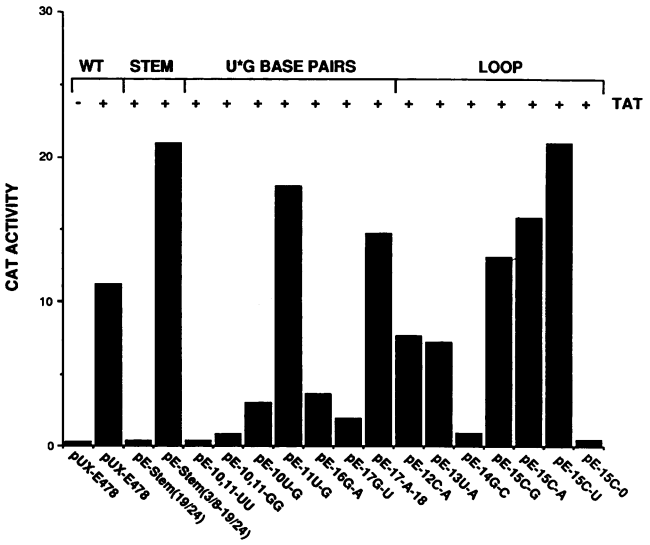

FIG. 3. CAT activities directed by mutated LTRs. Promoter activities of the various constructs were assayed by measuring CAT enzyme levels (11, 24). CAT activities are expressed as 10^3 cpm of [^{14}C]acetyl coenzyme A transferred to chloramphenicol per hour. Cell extracts were obtained from D17 cells after cotransfection with CAT plasmids in the presence (+) or absence (−) of the Tat expression plasmid pRS-Etat-M. Data are averages of at least five independent transfection assays.

with the uridine at position +10 to generate a TAR structure as shown in Fig. 2B; it is possible that the three-nucleotide loop shown is not stable and may actually be a five-base loop with a U · G pair at its base. This plasmid also showed higher levels of CAT activity than did the wild type (Fig. 3). The U · G base pair between positions +10 and +17 was abolished in plasmids pE-10U-G and pE-17G-U; the remaining U · G base pair is not adjacent to the helical stem and does not close the loop in these constructs. Both of these plasmids retained only limited activity (Fig. 2B and Fig. 3). Finally, in plasmid pE-16G-A, the U · G base pair that normally closes the loop was replaced with a U · A base pair. The resulting decrease in TAR activity to 32% of wild-type levels may be attributed to the location of the remaining U · G base pair within the helical stem. In sum, these data suggest that a U · G base pair at the base of the loop is an essential feature for EIAV TAR function. The precise nucleotide requirements at these positions and their role in Tat-mediated *trans* activation await the physical analysis of TAR structures and results of in vitro Tat-binding studies.

We next altered individual nucleotides within the remainder of the EIAV loop (Fig. 2C). The pyrimidines at positions +12(C) and +13(U) were each replaced with adenosine. The resultant constructs, pE-12C-A and pE-13U-A, showed *trans*-activation efficiencies of 69 and 64%, respectively, compared with the wild type. Substitution of the guanosine at position +14 by cytidine, pE-14G-C, resulted in a drastic reduction of *trans* activation to 8% of wild-type levels. Substitution of the cytidine at position +15 with a uridine, construct pE-15C-U, resulted in an increased TAR function, yielding 187% *trans*-activation efficiency. We then exchanged this cytidine with purines to generate pE-15C-A and pE-15C-G, which gave *trans*-activation efficiencies of 141 and 117%, respectively. An LTR in which this cytidine was deleted, pE-15C-O, failed to respond to Tat (Fig. 3). Our interpretation of these data is that position +15 must be occupied, by any base, to provide the proper spatial arrangement of adjacent bases or to stabilize the loop structure.

RNA analysis by primer extension. To determine whether the CAT activities directed by the EIAV TAR constructs described in the preceding section correlated with levels of CAT mRNA and to ensure that these mutations would not alter the transcription start site, CAT mRNA was examined by primer extension.

D17 cells were transfected with the mutated CAT plasmids in combination with EIAV Tat; in addition, all transfections included pRSV-CAT as an internal control. Total cytoplasmic RNA was prepared from transiently transfected cells and reverse transcribed with an oligonucleotide primer complementary to CAT sequences. The products of these reactions are predicted to be DNAs of 87 bp for pRSVCAT and 174 bp for the EIAV-CAT plasmids (plasmids pE-14G-C, pE-16G-A, and pE-10,11UU should yield products of about 206 bp due to additional polylinker sequences between the LTR and the *cat* gene).

An autoradiogram of the sequencing gel used to resolve the primer extension reaction products is shown in Fig. 4. The four lanes at the far left show dideoxy-sequencing reactions of pUX-E478 using the CAT primer. The next two lanes show the primer extension reaction products of CAT mRNAs from cells transfected with pUX-E478 and pRSPA (uninduced) or pRS-Etat-M. Tat expression caused a large increase in the amount of EIAV transcription, as seen by the uppermost band that corresponds to the primer extension product of CAT mRNA directed by EIAV LTR. The size of this band precisely identifies the EIAV transcription start site. A 150-nucleotide band whose size corresponds to

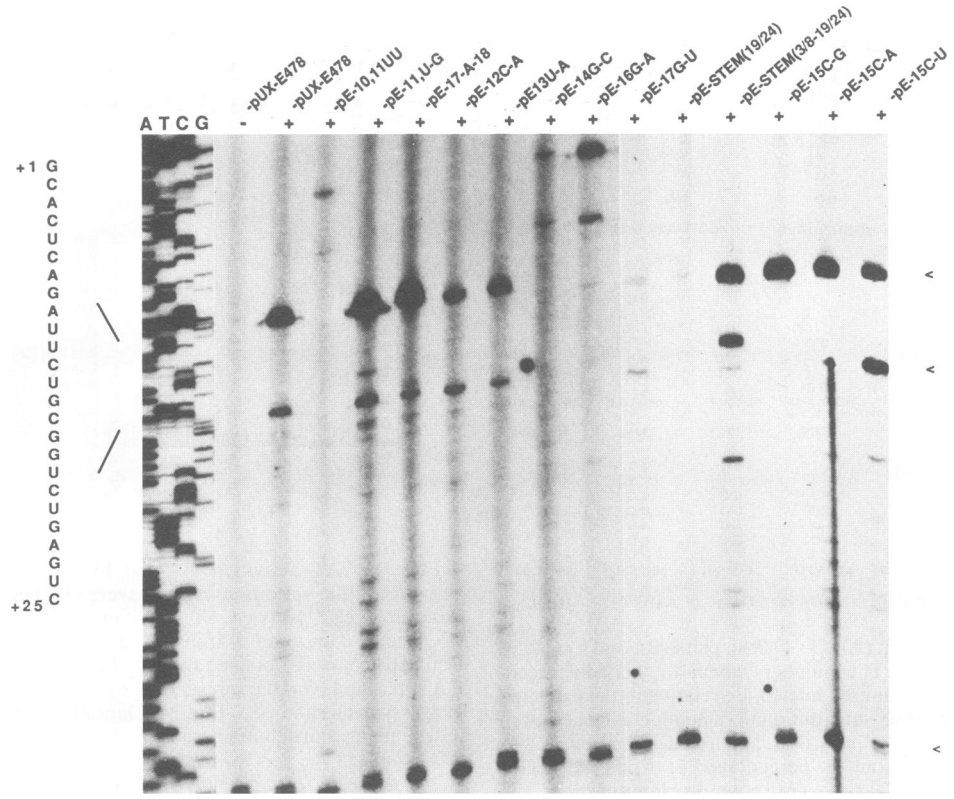

FIG. 4. Primer extension analysis of CAT mRNA produced in transfected cells. A radioactively labeled oligonucleotide complementary to the 5′ end of the *cat* gene was used for primer extension analysis on total cytoplasmic RNA from transfected D17 cells. Reaction products were resolved on a 6% polyacrylamide–7 M urea sequencing gel and visualized by autoradiography; a gap in the center of the autoradiogram has been removed, and the two sides were joined for this figure. D17 cells were cotransfected with pRSV-CAT and EIAV LTR CAT plasmids (indicated on top) in the absence (−) or presence (+) of a Tat expression plasmid. The same radiolabeled primer was used in dideoxynucleotide sequencing reactions on pUX-E478 (lanes A, T, C, and G) and run in parallel. The uppermost arrow at the far right indicates the major primer extension product with an expected size of 174 bases; the middle arrow indicates a band consistent with a pause in reverse transcription at position +25; the lower arrow indicates the product directed by the Rous sarcoma virus promoter used as an internal control. pE-10,11UU, pE-14G-C, and pE-16G-U have 32 nucleotides of polylinker between the LTR and *cat* and thus produce an expected fragment of 206 bases. The printed sequence at the extreme left corresponds to the portion of the gel that spans the RNA start site (+1) and the end of the double-stranded stem (+25).

termination of primer elongation at the base of the stem (position +25) was also observed in most reactions and may result from reverse transcriptase pausing at this double-stranded region. This provides physical evidence in support of the proposed stem structure. The lowest band present in all lanes represents the CAT mRNA transcript from the Rous sarcoma virus promoter used as an internal control of transfection efficiency, RNA isolation, and primer extension.

The remaining lanes show the primer extension products of CAT mRNA directed by the mutated LTRs and demonstrate that mutations introduced into EIAV TAR affected the steady-state levels of CAT mRNA. The increased CAT activities observed with pE-15C-U, pE-15C-G, pE-15C-A, pE-17-A-18, and pE-11U-G (Fig. 3) correlated with increased amounts of CAT mRNA (Fig. 4). Similarly, pE-12C-A, pE-13U-A, pE-16G-A, and pE-17G-U showed reduced CAT activities (Fig. 3) and proportional reductions in CAT mRNA (Fig. 4). Alterations at position +14 (pE-14G-C), deletion of the two uridines present at the beginning of the loop (pE-10,11UU), and disruption of the stem (pE-Stem+19/+24) were shown to be deleterious by CAT enzyme assays (Fig. 3). These plasmids were capable of supporting levels of transcription that were only slightly higher than the basal level (Fig. 4). These data indicate that the levels of mRNA directed by EIAV LTRs were proportional to the observed CAT enzyme activities and that EIAV Tat acts through TAR to increase steady-state levels of viral RNA.

DISCUSSION

The EIAV LTR region between −31 and +22 (with respect to the RNA start site) was shown previously to contain the *cis*-acting TAR (11). Moreover, we have observed that a composite LTR, containing the HIV-1 U3 region joined at the RNA start site to the EIAV R region, was specifically *trans* activated by EIAV Tat (4), indicating that the EIAV LTR region from +1 to +22 must contain a functional EIAV TAR. This region contains two inverted repeats, and hence the RNA transcript can assume a hairpin structure (Fig. 1). It has been shown that HIV-1 and other primate lentiviruses possess similar secondary structures at the 5′ ends of their RNAs which function as TARs (2, 13, 14, 17, 18, 32). In EIAV RNA, this proposed structure has a shorter stem, lacks the uridine bulge, and has a loop with

four nucleotides rather than the six observed in the primate lentivirus TAR elements.

Primer extension analysis identified the position of the EIAV RNA start site and revealed that all mutated LTRs initiated transcription at this site. Furthermore, a second band was observed in the primer extension reactions that is consistent with premature termination of reverse transcription at the putative stem, thus providing physical evidence for the proposed hairpin structure. CAT enzyme levels were proportional to mRNA levels, and Tat-mediated *trans* activation was accompanied by increases in steady-state levels of mRNA.

The biological importance of the EIAV stem structure was established by the observations that clustered mutations that disrupted this structure abolished the Tat-mediated *trans* activation and compensatory mutations that restored the base pairing also recovered the response to Tat. Previous deletion analysis of the EIAV LTR revealed that a construct lacking sequences from position +23 to the 3′ end of the LTR was still *trans* activated (11). This LTR would produce only a 4-bp stem, suggesting that the secondary structure in the upper portion of EIAV stem makes the major contribution to *trans* activation. This observation is similar to that reported for HIV-1, in which disruption of base pairing in the upper portion of the HIV-1 stem was shown to abolish responsiveness to Tat, but LTRs altered downstream of position +45 were fully Tat inducible (14, 17, 32). Thus, the biological function of the EIAV stem seems to rely on a stretch of 4 bp, making EIAV TAR the smallest TAR yet reported for lentiviruses. Interestingly this 4-bp stem is homologous to the 4 bp present downstream of the HIV-1 bulge and comprises the minimal lower portion of the stem required for Tat activation. Construct pE-Stem(+3/+8−+19/+24) presents a stem with an altered primary sequence and showed a Tat response that was greater than wild type. Curiously, HIV-1 TARs with altered primary sequences, but not secondary structures, in the 4-bp stretch cited above exhibited different levels of Tat responsiveness (14, 32). Most of the alterations decreased TAR function (14, 32), but one that contained a mismatched base pair showed 180% of wild-type activity (14). In addition, the in vitro binding affinity of HIV-1 Tat protein for a TAR RNA fragment containing an altered sequence in this stretch was reduced (29). While it is apparent that the secondary structure of this part of the stem is critical, the role of the primary sequence is less clear. The conservation of these bases downstream of the U · G base pairs in EIAV or a bulge in HIV-1 may reflect a significant common role in Tat activation. A detailed mutagenesis analysis in this region combined with in vitro binding studies should reveal its contribution to Tat activation.

We have represented the EIAV TAR structure as a four-base loop that is closed by two U · G base pairs. Thermodynamic analyses have indicated that a U · G pair is about as stable as a U · A pair in the A-form helix of RNA (25). However, the geometry of the U · G pair is different than that of a U · A pair and changes the local structure of the RNA helix; in the non-Watson-Crick base pair, the purine and pyrimidine are shifted toward the minor and major grooves, respectively (25). The mutagenesis studies presented here indicated that the U · G pairs that close the EIAV TAR loop are essential for *trans* activation and that a single pair at the base of the loop is sufficient. EIAV TAR lacks the bulge structure that is present in the stem of HIV-1 TAR; the unpaired uridines in this bulge provide a unique three-dimensional structure that is bound by the HIV-1 Tat protein (3, 8, 29, 30, 39). We suggest that the non-Watson-Crick, U · G pairs in EIAV TAR are the functional counterpart of the bulge structure in HIV-1 TAR. The U · G pairs adjacent to a loop may provide a distinctive conformation that is recognized by EIAV Tat. By analogy, the helical irregularity associated with a U · G pair in the acceptor stem of $tRNA^{Ala}$ was shown to confer acceptor identity by its cognate aminoacylating enzyme (22).

The EIAV TAR loop sequence (positions +12 to +17) consists of the sequence CUGC. The individual nucleotide substitutions of this sequence (Fig. 2C) demonstrated that the relative contribution of each is not uniform for Tat activation. Mutations in the cytidine and uridine at positions +12 and +13, respectively, showed modest effects on TAR function. The +14 guanosine residue is the most critical position, exhibiting a deleterious phenotype when altered. The cytidine at position +15 is unique; all of the substitutions at this position resulted in increased levels of gene expression. The magnitude of the effect was dependent on the identity of the replacement base; substitution of cytidine with uridine resulted in the highest *trans* activation (187% of the wild-type level). Deletion of this cytidine abolished activity, indicating the structural importance of this position. It is unclear at present whether the alterations of loop nucleotides affected TAR function by influencing loop structure and stability or because direct contacts with Tat were altered. While it has been suggested that HIV-1 TAR may contain binding sites for both Tat and cellular proteins (8, 15, 16, 21, 29), it would appear that the small size of EIAV TAR would preclude binding by other proteins in addition to Tat.

We conclude that the secondary structure of the stem, the presence of at least one U · G base pair at the base of the loop, and the specific nucleotide identity of the loop comprise a complex set of requirements for efficient EIAV Tat-mediated *trans* activation. Primate lentivirus and EIAV Tat proteins share structural similarities and utilize an RNA TAR located downstream of the transcription initiation site. In addition, EIAV and HIV-1 TARs were shown to be structurally and functionally similar. Thus, we suggest that these two systems have similar mechanisms of Tat function. Although specific proteins are involved in the Tat-mediated *trans* activation in each system, the relative simplicity of EIAV may facilitate the understanding of the basic mechanism.

ACKNOWLEDGMENTS

We thank Laraine Rudy for secretarial help. We are gratefully to L. Martarano and R. Carroll for assistance.

M.C. is supported by CNPq fellowship 202395/89.8, Conselho Nacional Desenvolvimento e Tecnologia, Brazil.

REFERENCES

1. **Arya, S. K., and R. C. Gallo.** 1988. Human immunodeficiency virus type 2 long terminal repeat: analysis of regulatory elements. Proc. Natl. Acad. Sci. USA **85**:9753–9757.
2. **Berkhout, B., and K.-T. Jeang.** 1989. *trans* activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the *trans*-acting responsive hairpin: a quantitative analysis. J. Virol. **63**:5501–5504.
3. **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.
4. **Carroll, R., L. Martarano, and D. Derse.** 1991. Identification of lentivirus Tat functional domains through generation of equine infectious anemia virus/human immunodeficiency virus type 1 *tat* gene chimeras. J. Virol. **65**:3460–3467.
5. **Cheevers, W. P., and T. C. McGuire.** 1985. Equine infectious

anemia virus: immunopathogenesis and peristence. Rev. Infect. Dis. 7:83–88.
6. **Cullen, B. R.** 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell **46**:973–982.
7. **Cullen, B. R., and W. C. Greene.** 1990. Functions of the auxiliary gene products of the human immunodeficiency virus type 1. Virology **178**:1–5.
8. **Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner.** 1990. HIV-1 Tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. EMBO J. **9**:4145–4153.
9. **Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, M. Skinner, and R. Valerio.** 1989. Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. Proc. Natl. Acad. Sci. USA **86**:6925–6929.
10. **Dorn, P., L. DaSilva, L. Martarano, and D. Derse.** 1990. Equine infectious anemia virus *tat*: insights into the structure, function, and evolution of lentivirus *trans*-activator proteins. J. Virol. **64**:1616–1624.
11. **Dorn, P. L., and D. Derse.** 1988. *cis*- and *trans*-acting regulation of gene expression of equine infectious anemia virus. J. Virol. **62**:3522–3526.
12. **Endo, S.-I., S. Kubota, H. Siomi, A. Adachi, S. Oroszlan, M. Maki, and M. Hatanaka.** 1989. A region of basic amino-acid cluster in HIV-1 tat protein is essential for trans-acting activity and nucleolar localization. Virus Genes **3**:99–110.
13. **Feng, S., and E. C. Holland.** 1988. HIV-1 tat trans-activation requires the loop sequence within tar. Nature (London) **334**: 165–168.
14. **Garcia, J. A., D. Harrich, E. Soultanakis, F. Wu, R. Mitsuyasu, and R. B. Gaynor.** 1989. Human immunodeficiency virus type I LTR TATA and TAR region sequences required for transcriptional regulation. EMBO J. **8**:765–778.
15. **Gatignol, A., A. Kumar, A. Rabson, and K.-T. Jeang.** 1989. Identification of cellular proteins that bind to the human immunodeficiency virus type 1 trans-activation-responsive TAR element RNA. Proc. Natl. Acad. Sci. USA **86**:7828–7832.
16. **Gaynor, R., E. Soultanakis, M. Kuwabara, J. Garcia, and D. S. Sigman.** 1989. Specific binding of a HeLa cell nuclear protein to RNA sequences in the human immunodeficiency virus transactivating region. Proc. Natl. Acad. Sci. USA **86**:4858–4862.
17. **Hauber, J., and B. R. Cullen.** 1988. Mutational analysis of the *trans*-activation-responsive region of the human immunodeficiency virus type I long terminal repeat. J. Virol. **62**:673–679.
18. **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 an HIV *trans* activator. Mol. Cell. Biol. **8**:2555–2561.
19. **Kobayashi, K., and Y. Kono.** 1967. Serial passages of equine infectious anemia virus in horse leucocyte cultures. Natl. Inst. Anim. Health Q. (Tokyo) **7**:1–7.
20. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
21. **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. USA **87**:3624–3628.
22. **McClain, W. H., Y-M. Chen, K. Foss, and J. Schneider.** 1988. Association of transfer RNA acceptor identity with a helical irregularity. Science **242**:1681–1684.
23. **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.
24. **Neumann, J. R., C. A. Morency, and K. O. Russian.** 1987. A novel rapid assay for chloramphenicol acetyltransferase gene expression. BioTechniques **5**:444–447.
25. **Puglisi, J. D., J. R. Wyatt, and I. Tinoco.** 1990. Solution conformation of an RNA hairpin loop. Biochemistry **29**:4215–4226.
26. **Rappaport, J., S.-J. Lee, K. Khalili, and F. Wong-Staal.** 1989. The acidic amino terminal region of HIV-1 Tat protein constitutes an essential activating domain. New Biol. **1**:101–110.
27. **Rosen, C. A.** 1991. Regulation of HIV gene expression by RNA-protein interactions. Trends Genet. **7**:9–14.
28. **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.
29. **Roy, S., U. Delling, C.-H. Chen, C. A. Rosen, and N. Sonenberg.** 1990. A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. Genes Dev. **4**:1365–1373.
30. **Roy, S., N. 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.
31. **Rushlow, K., K. Olsen, G. Stiegler, S. L. Payne, R. C. Montelaro, and C. Issel.** 1986. Lentivirus genomic organization: the complete nucleotide sequence of the env gene region of equine infectious anemia virus. Virology **155**:309–321.
32. **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.
33. **Selby, M. J., and B. M. Peterlin.** 1990. Trans-activation by HIV-1 Tat via a heterologous RNA binding protein. Cell **62**:769–776.
34. **Siomi, H., H. Shida, M. Maki, and M. Hatanaka.** 1990. Effects of a highly basic region of human immunodeficiency virus *tat* protein on nucleolar localization. J. Virol. **64**:1803–1807.
35. **Sodroski, J., R. Patarca, C. A. Rosen, F. Wong-Staal, and W. A. Haseltine.** 1985. Localization of the trans-activating region on the genome of human T-cell lymphotropic virus type III. Science **229**:74–77.
36. **Stephens, R. M., D. Derse, and N. R. Rice.** 1990. Cloning and characterization of cDNAs encoding equine infectious anemia virus Tat and putative Rev proteins. J. Virol. **64**:3716–3725.
37. **Subramanian, T., M. Kuppuswamy, L. Venkatesh, A. Srinivasan, and G. Chinnadurai.** 1990. Functional substitution of the basic domain of the HIV-1 trans-activator tat, with the basic domain of the functionally heterologous Rev. Virology **176**:178–183.
38. **Walker, M. D., T. Edlund, A. M. Boulet, and W. J. Rutter.** 1983. Cell-specific expression controlled by the 5′-flanking region of insulin and chymotrypsin genes. Nature (London) **306**:557–561.
39. **Weeks, K. M., C. Ampe, S. C. Schultz, T. A. Steitz, and D. M. Crothers.** 1990. Fragments of the HIV-1 Tat protein specifically bind TAR RNA. Science **249**:1281–1285.