Sequences Downstream of the RNA Initiation Site Regulate Human T-Cell Lymphotropic Virus Type I Basal Gene Expression

FATAH KASHANCHI, JANET F. DUVAL, PAUL F. LINDHOLM, MICHAEL F. RADONOVICH,
AND JOHN N. BRADY*

Laboratory of Molecular Virology, National Cancer Institute, Bethesda, Maryland 20892

Received 11 September 1992/Accepted 22 January 1993

Sequences which control basal human T-cell lymphotropic virus type I (HTLV-I) transcription probably play an important role in initiation and maintenance of virus replication. We have identified and analyzed a 45-nucleotide sequence (downstream regulatory element 1 [DRE 1]) at the boundary of the R/U5 region of the long terminal repeat which is required for HTLV-I basal transcription. The basal promoter strength of constructs that contained deletions in the R/U5 region of the HTLV-I long terminal repeat were analyzed by chloramphenicol acetyltransferase assays following transfection of Jurkat T cells. We consistently observed a 10-fold decrease in basal promoter activity when sequences between +202 to +246 were deleted. By reverse transcriptase polymerase chain reaction RNA analysis, we confirmed that the drop in chloramphenicol acetyltransferase activity was paralleled by a decrease in the level of steady-state RNA. DRE 1 did not affect the level of Tax₁ transactivation. Using a gel shift assay, we have purified a highly enriched fraction that could specifically bind DRE 1. This DNA affinity column fraction contained four detectable proteins on sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis: p37, p50, p60, and p100. The affinity column fraction stimulated HTLV-I transcription approximately 12-fold *in vitro*. No effect was observed with the human immunodeficiency virus or adenovirus major late promoters. Following renaturation of the proteins isolated from an SDS-containing gel, p37, but not the other protein fractions, was able to specifically bind to DRE 1.

Replication of human T-cell lymphotropic virus type I (HTLV-I) is regulated at the transcriptional and posttranscriptional levels by viral gene products, p40^{tax} (Tax₁) and p27^{rex1} (Rex₁) (1, 11, 13, 15, 17, 42). The Tax₁ transactivator positively regulates expression of viral mRNAs and several cellular growth regulatory genes and cytokines including *c-sis*, *c-fos*, interleukin-2 (IL-2), IL-2Ra, granulocyte-macrophage colony-stimulating factor, vimentin, proenkephalin, and PTHrP (12, 21, 22, 28, 36, 40, 44, 46). The Rex₁ gene product regulates the expression and transport of singly-spliced and unspliced viral mRNAs which encode the structural gene products Gag, Env, and Pol. Recently, two new viral mRNAs which code for two previously unidentified proteins, Tof and Rof, have been identified (8). The function of these proteins remains to be established.

The 5'-U3 region of the HTLV long terminal repeat (LTR) contains important elements needed for Tax₁ transactivation. The 21-bp repeat elements (TRE-1) confer Tax₁ responsiveness and function in either orientation, and two or more of these elements confer Tax₁ responsiveness to heterologous promoters (5, 30, 35, 39). A second Tax₁-responsive element (TRE-2) is located between the two proximal 21-bp repeats at -117 to -163 and is a binding site for SP1, TIF-1, Ets, and Myb (3, 4, 14, 26, 32). Sequences downstream of the RNA initiation site, encompassing the R region and the 5' portion of the U5 region, have also been shown to be important for HTLV-I gene expression. Nakamura et al. reported that a 136-bp fragment (+104 to +240) increased the level of HTLV-I gene expression (31). Consistent with this report, Seiki et al. reported that sequences between +32 and +266 elevated mRNA synthesis (38). The downstream

element was orientation independent but was effective only when it was located between the transcription initiation site and the translation initiator site. The downstream regulatory sequence was not Tax₁ responsive. A distinct regulatory element, described by Seiki et al. (38), which may function as an RNA element, is located between +266 and +347.

Downstream regulatory elements have been shown to play an important role in the positive and negative regulation of several viral and cellular genes (7, 9, 16, 20, 25, 27, 29, 34, 43, 45, 47). For example, in the adenovirus (Ad) type 5 E1A promoter a 10-fold reduction in the steady-state RNA level is observed in response to a single-base deletion 399 nucleotides downstream of the transcription start site (34). In the bovine leukemia virus promoter, deletion of a positive control element in a 250-bp segment downstream of the RNA start site reduces gene expression by 87% (9). When this 250-bp position-dependent sequence was inserted downstream of the transcription start site of a simian virus 40 early promoter, chloramphenicol acetyltransferase (CAT) activity increased 60-fold. In contrast to the upstream bovine leukemia virus enhancer, the 250-bp downstream element is active in both bovine leukemia virus-infected and uninfected cells, suggesting a role for cellular factors (9). In the *c-myc* gene, a downstream transcription modulator is required for transcription from the *c-myc* P2 promoter. The modulator is orientation dependent and promoter specific, increasing transcription from the *c-myc* P2 promoter and simian virus 40 upstream early promoter (pSVE2) but not the *c-myc* P1 promoter or the simian virus 40 downstream early promoter (pSVE1) (47). In minute virus of mice, a downstream 365-bp element (DPE) is essential for basal transcription from the P38 promoter (20).

In the work described in this report, we have analyzed the effect of the R/U5 region and characterized 45 bp that is

* Corresponding author.

needed for Tax₁-independent basal transcription. Using a gel shift assay, we have purified a highly enriched protein fraction that could specifically bind DRE 1. This DNA affinity column fraction contained four proteins detectable on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE): p37, p50, p60, and p100. This protein fraction stimulated HTLV-I transcription approximately 12-fold in vitro. One of these proteins, p37, was able to specifically bind to DRE 1.

MATERIALS AND METHODS

Plasmid constructs. Primers were synthesized with 5' *Bgl*II-*Xho*I and 3' *Bgl*III sites. After polymerase chain reaction (PCR) amplification, the DNA was cleaved with *Bgl*II, purified on a 1.5% agarose gel, and electroeluted in 1× TAE buffer at 50 V for 2 to 3 h. After ethanol precipitation, the DNA was ligated into the unique *Bgl*III site (0/4380) in pCAT3M. Each construct contains a single *Xho*I site for analysis and cloning. The positive numbers (dl+15, dl+42, dl+70, dl+102, dl+157, dl+202, dl+220, dl+246, and dl+262) indicate the length of DNA included in each construct downstream of the +1 start site.

Transfections and CAT assays. Jurkat T cells were harvested at the mid- to late log phase, washed once in phosphate-buffered saline (PBS) without Mg²⁺ or Ca²⁺, pelleted, and resuspended in RPMI 1640 at a concentration of 10⁷ cells per 250 µl. Then 20 µg of each plasmid DNA construct and 250 µl of cells were mixed, transferred to electroporation vials, and electroporated at 250 V and a capacitance of 800 µF. The transfected cells were incubated for 10 min on ice, plated out in 10-cm dishes containing 10 ml of RPMI 1640 plus 10% FCS, 100 U of penicillin per ml, 100 µg of streptomycin per ml, and 2 mM L-glutamine, and incubated at 37°C. At 24 h later, cells were harvested, washed once in PBS (without Ca²⁺ or Mg²⁺), and resuspended in 150 µl of 0.25 M Tris (pH 7.8). After three freeze-thaw cycles and centrifugation for 5 min at 4,000 rpm (DuPont Sorvall RT6000 tabletop centrifuge), the 150 µl of supernatant was transferred to Eppendorf tubes and heated for 5 min at 65°C. The supernatant was centrifuged again for 5 min and transferred to a new Eppendorf tube. CAT assays were then performed with approximately 30 µg of protein.

Jurkat whole-cell and nuclear extracts. For whole-cell extract preparations, we used a modification of the procedure of Manley et al. (24). The following steps were used. (i) Jurkat T cells were harvested at 1,000 rpm (DuPont Sorvall RT6000 tabletop centrifuge) and 4°C for 10 min in culture media. (ii) After harvesting, the pellet was resuspended in 0.5 volume of cold PBS, mixed well, counted, and pelleted as above. (iii) The cells were then washed twice in 0.1 volume of cold buffer A (10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; pH 7.9] at 4°C, 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM dithiothreitol [DTT]) and pelleted as above. (iv) After being centrifuged, the washed cell pellet was resuspended in 20 µl of cold buffer C (20 mM HEPES [pH 7.9], 25% [vol/vol] glycerol, 0.42 M NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride, [added fresh, stock in isopropanol], 0.5 mM DTT [added fresh], 0.1% Nonidet P-40) per 10⁷ cells, incubated for 10 min on ice, vortexed for 10 s, and centrifuged at 10,000 rpm for 10 min in a microcentrifuge. (v) The lysed cell supernatant was retained, diluted with 80 µl of cold modified buffer D (20 mM HEPES [pH 7.9], 20% [vol/vol] glycerol, 0.05 M KCl, 0.2 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride, 0.5 mM DTT) per 10⁷ cells, and stored at -70°C as

100-µl aliquots. Protein concentrations were determined prior to storage.

Nuclear extracts were prepared by following steps (i) through (iii) above. The washed cell pellets were then resuspended in 20 µl of cold buffer A-0.1% Nonidet P-40 per 10⁷ cells, incubated on ice for 10 min, mixed briefly (10 s) by vortexing, and centrifuged at 10,000 rpm at 4°C for 10 min in a microcentrifuge. The supernatant was carefully removed and discarded, and the nuclear pellet was resuspended in 15 µl of cold buffer C without Nonidet P-40 per 10⁷ cells. After a 15-min incubation on ice with intermittent vortexing (5 s), the mixture was centrifuged in a microcentrifuge as above. The nuclear supernatant was retained, diluted with 75 µl of cold modified buffer D per 10⁷ cells, and stored at -70°C as 100-µl aliquots. Protein concentrations were determined prior to storage.

Fractionation of nuclear extracts. Jurkat T cells (approximately 3 × 10¹⁰) at the mid- to late log phase of growth were used as startup material for purification steps. We used modified buffer D to either equilibrate columns or dialyze the proteins after chromatography. Band shift analysis was used as the detection method to monitor the protein of interest during purification steps. All chromatography steps were performed at 4°C with a flow rate of approximately 0.5 ml/min. Nuclear extracts were prepared by the procedure indicated in the previous section. Approximately 288 mg of protein was recovered from the nuclear extract and was loaded onto a Sephadex G-25 column (bed volume, 10 ml, medium; Pharmacia), and a total of 89 mg was recovered from the flowthrough. Next, the recovered extract was loaded onto a Sephacryl S-200 sizing column (bed volume, 40 ml, high resolution, Pharmacia). We recovered a total of 35 mg, which was loaded onto a fast protein liquid chromatography (FPLC) Mono-S ion-exchange column (bed volume, 10 ml; Pharmacia), and the flowthrough which contained the activity (15 mg total) was loaded onto an FPLC-DNA affinity column (18). The DNA affinity column was prepared by using CNBr-activated Sepharose CL-2B (2 ml of resin) attached to 200 µg of gel-purified double-stranded oligonucleotide (DNA sequence of +188 to +262 of the HTLV-I R/U5 region). A 40-ml linear gradient of KCl (0.05 to 1 M) in buffer D was used to elute 1-ml fractions. The fractions were then concentrated to approximately 0.1 ml in a microconcentrator (Centricon 10; Amicon) and desalted on a Sephadex G-25 spin column (Pharmacia). Highly enriched fractions were assayed for band shift activity, and fractions 13 to 18 (elution at approximately 0.2 to 0.25 M KCl) were dialyzed against buffer D and used for in vitro transcription analysis (2). A total of 19 µg of protein was recovered from DNA affinity column pooled fractions 13 to 18.

Gel retardation analysis. The oligonucleotide probe was labeled with [γ -³²P]ATP by using T4 kinase enzyme. A 12.5-ng portion of the labeled oligonucleotide was incubated with 7.5 µg of Jurkat whole-cell extract and 3 µg of double-stranded poly(dI-dC)-poly(dI-dC) in gel shift binding buffer (10 mM Tris [pH 7.5], 40 mM NaCl, 1 mM DTT, 1 mM EDTA) for 20 min at 24°C. For competition studies, 300 or 600 ng of unlabeled oligonucleotide was added to the incubation mixture. After incubation, DNA-protein complexes were analyzed on a 4% native acrylamide gel with 0.25× TBE as the running buffer.

Protein renaturation. Pooled fractions 13 to 18 from the DNA affinity chromatography column were separated by preparative SDS-PAGE (10% polyacrylamide separating gel), and the proteins were visualized by using 0.3 M CuCl₂ (26). Proteins of various molecular weights were excised

from the gel and destained with three 10-min changes of a buffer containing 0.25 M Tris hydrochloride (pH 9) and 0.25 M EDTA. Proteins were passively eluted from the gel slices in 2 ml of buffer containing 50 mM HEPES (pH 7.9), 0.1 mM EDTA, 0.1% SDS, 5 mM DTT, and 150 mM NaCl by incubation overnight at room temperature with gentle rotation. The eluate was precipitated with 4 volumes of acetone (-20°C) and centrifuged for 30 min at 10,000 rpm in an HB-4 rotor. The pelleted precipitate was washed with 80% acetone (-20°C) and centrifuged again to collect the precipitate. The pellet was resuspended in 200 μl of X50 buffer (20 mM HEPES [pH 7.9], 20% glycerol, 50 mM KCl, 1 mM EDTA, 1 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride) containing 6 M guanidine hydrochloride. The samples were dialyzed over a 48-h period at 4°C against a 2,000-fold excess of X50 buffer. Renatured proteins were stored at -70°C for subsequent analyses.

Reverse transcription and PCR analysis. Total-cell RNA was harvested by lysis with 4 M guanidinium isothiocyanate, and the lysate was centrifuged for 18 h at 36,000 rpm in a Beckman SW55Ti rotor at 18°C . The RNA was harvested, ethanol precipitated, and quantitated by measurement of A_{260} . Aliquots of cellular RNA were then reverse transcribed with 3' antisense primers specific for either the HTLV-I CAT or cellular actin gene. For HTLV-I RNA analysis, a 15-nucleotide primer, 5'-CTGGATATTACGGCC-3', located immediately adjacent to the *Pvu*II site in the CAT gene was utilized. The reverse transcription reaction was performed with 400 mM deoxynucleotides and 400 U of mouse mammary tumor virus reverse transcriptase (RT) for 1 hr at 37°C as previously described (44). The reverse-transcribed products were then subjected to PCR in the presence of 1 pmol of 5' sense and 3' antisense primers, 200 mM deoxynucleoside triphosphates, and 2.5 U of *Taq* polymerase in a buffer containing 50 mM KCl and 1.5 mM MgCl_2 . Samples were subjected to 40 cycles of amplification consisting of denaturation for 1 min at 94°C , primer annealing for 1 min at 53°C , and polymerization for 2 min at 72°C . Two sets of primers were used in the same test tube for each PCR amplification. One set of primers amplified HTLV-I-specific mRNA, and the other set amplified actin-specific mRNA. First-strand cDNA synthesis was performed with 5'-CTG GATATTACGGCC-3' primer at the *Pvu*II site of the CAT gene, and PCR amplification was performed with +1 to +15 of HTLV-I sequence 5'-GGCTCGCATCTCTCC-3'. The use of a single set of primers to amplify all HTLV-I RT samples eliminates the hybridization variability and efficiency of multiple primers. An actin-specific sequence served as an internal test control and was defined by 5'-TTCTACAAT GAGCTGCGTGT-3' and 5'-GCCAGACAGCACTGTGT TGG-3' sense and antisense primers, respectively. The amplified 636-bp actin signal was detected by an internal 40-base probe, 5'-ACTACCTCATGAAGATCCTCACCGA GCGCGCTACAGCTT-3'. Aliquots (50 μl) of the PCR products were loaded onto 1.2% agarose gels and electrophoresed. The gels were then soaked in 1 M NaOH-1.5 M NaCl for 15 min and neutralized for 30 min prior to Southern blotting. The gels were blotted onto nitrocellulose overnight, dried, baked, prehybridized, and probed with either CAT or actin probes (approximately 300 ng of ^{32}P -end-labeled probes).

For RNA 5'-end analysis, primer extension reactions were performed with a [γ - ^{32}P]ATP-end-labeled antisense primer (approximately 2×10^6 cpm), 5'-TGTAACGGCGCAGAAC-3', positioned 263 bases downstream of the RNA initiation site. Then 5 μg of RNA, isolated from transfected or

untransfected Jurkat T cells, was added to each reaction. The conditions for the reverse transcription reaction are described above. Products were phenol-chloroform extracted and analyzed on a denaturing acylamide-urea gel next to a sequence reaction from either M13mp18 or the HTLV-I CAT plasmid.

In vitro transcription: unfractionated extracts. The in vitro transcription buffer contained 10 mM HEPES (pH 7.9), 50 mM KCl, 0.5 mM EDTA, 1.5 mM DTT, 6.25 mM MgCl_2 , and 8.5% glycerol. The DNA template pU3R-CAT was linearized with *Hind*III and added to a concentration of 13.2 $\mu\text{g}/\text{ml}$ (200 ng per reaction). pAdML was linearized with *Bam*HI and added at a concentration of 13.2 $\mu\text{g}/\text{ml}$ (200 ng per reaction). HeLa whole-cell extracts, prepared as described previously (24), were added to a final concentration of 2.4 mg/ml (40 μg per reaction). Oligonucleotide affinity fraction 16 was titrated over a concentration range of 160 to 1,120 ng per reaction. Nucleoside triphosphates in water were added to a final concentration of 500 μM unless otherwise stated. Transcription reactions were terminated by the addition of 20 mM Tris-HCl (pH 7.8), 150 mM NaCl, and 0.2% SDS. The quenched reactions were extracted with phenol-chloroform, chloroform and precipitated with 2.5 volumes of ethanol and 0.1 volume of 3.0 M sodium acetate. Following centrifugation, RNA pellets were resuspended in 12 μl of formamide denaturation mix containing xylene cyanol and bromophenol blue, heated at 90°C for 3 min, and electrophoresed at 400 V in a 4% polyacrylamide gel (acrylamide/bisacrylamide ratio, 19:1) containing 7 M urea (pre-run at 200 V for 30 min) and $1\times$ TBE. Gels were exposed to Kodak X-Omat XR-5 film at -70°C with intensifying screens for autoradiography.

RESULTS

R/U5 region of HTLV-I is important for basal transcription.

We have constructed a series of 3' deletion mutants and inserted them upstream of the CAT gene in pCAT3M. This was done by using an identical 5' primer and different 3' primers to amplify the fragment of interest by PCR (see Materials and Methods). The resulting nine constructs contained intact U3 sequences and a varying number of bases at the R and U5 regions of the HTLV-I LTR (Fig. 1A). After the fragments were subcloned into the CAT vector, the plasmids were electroporated into Jurkat T cells at the mid-to late log phase of growth and CAT assays were performed with equal amounts of protein lysates (30 μg) (Fig. 1B). The wild-type construct (-450 to $+262$) is functional in Jurkat T cells and can be transactivated by HTLV-I Tax₁ (Fig. 1B, lanes 1 and 2). In contrast, constructs that include sequences from -450 to $+202$ have very little basal activity (lanes 3 to 8). The basal transcription was increased 10- to 20-fold when downstream sequences $+203$ to $+262$ were present in the HTLV-I promoter (lanes 9 to 11). The R/U5 region regulatory element, required for basal transcription, was named downstream regulatory element 1 (DRE 1). This basal transcription was independent of Tax₁ transactivation (data not shown).

We next analyzed mRNA expression in cells transfected with the 3' deletion constructs. RNA was isolated and subjected to RT and PCR amplification. Two sets of primers were used in each reaction, one set to amplify HTLV-I-specific sequences and the second to amplify actin-specific mRNA as an internal control. It is important to note that a single set of PCR primers was used to amplify all HTLV-I RT samples to avoid quantitative hybridization variability in

FIG. 1. Downstream deletion mutants of the HTLV-I promoter. (A) 3' deletion constructs of the HTLV-I promoter were made by PCR amplification. Subsequently, the DNA fragments were cloned into pCAT3M. All constructs contained intact U3 sequences and various sequences of the R or R/U5 region. (B) CAT assays of 3' HTLV-I deletion mutants in Jurkat T cells. Lane 1 represents a positive control with the intact HTLV-I promoter cotransfected with Tax₁ expression vector. All other transfections were done in the absence of Tax₁. CAT assays were performed as described in Materials and Methods.

the PCR. The PCRs were divided in half, run on agarose gels, Southern blotted, and hybridized with either CAT or actin probes. Consistent with the results of the CAT assay, RNA expression levels of plasmids pds+220 (Fig. 2A, lane 7), pds+246 (lane 8), and pds+262 (lane 9) were increased compared with those of deletion mutants pds+157 (lane 5) and pds+202 (lane 6). As a control, the levels of actin mRNA in these RT-PCRs was measured, and similar amounts of actin-amplified mRNA were observed (Fig. 2B).

For RNA 5'-end analysis, primer extension reactions were performed with a [γ -³²P]ATP-end-labeled antisense primer (approximately 2×10^6 cpm), 5'-TGTAACGGCGCAGAAC-3', which includes HTLV-I sequences +248 to +263 downstream of the RNA initiation site. Then 5 μ g of RNA, isolated from transfected or control untransfected Jurkat T cells, was added to each primer extension reaction. In a parallel reaction, DNA dideoxy sequence reactions were performed on a single-stranded M13mp18 DNA or a denatured double-stranded HTLV-I DNA template. In the

HTLV-I DNA dideoxy sequence reaction, the same +248 to +263 oligonucleotide used for primer extension of RNA was used. Electrophoretic analysis of the primer extension products resulted in the appearance of one major DNA band (Fig. 2C, lane 1, and D, lane 2). Analysis of the primer extension products demonstrated that the major product was 263 bases in length and corresponded to the authentic RNA initiation site previously identified for HTLV-I (Fig. 2C and D). The minor primer extension products of 255 and 258 bases, as well shorter products, were not reproducibly detected in the primer extension analysis and most probably reflect incomplete extension products. No primer extension product was detected following incubation with the RNA from control Jurkat cells (Fig. 2C, lane 2; and D, lane 1), demonstrating the specificity of the primer for HTLV-I RNA.

Jurkat nuclear extracts contain a protein(s) which specifically interacts with DRE 1. In light of the requirement for downstream sequences for HTLV-I basal transcription, we were interested in identifying a transcriptional regulatory protein(s) which interacted specifically with this control region. An oligonucleotide was synthesized by using HTLV-I sequences +188 to +262 to analyze the DRE 1-binding protein(s) via a band shift assay. Using Jurkat nuclear extract, we demonstrated protein binding to oligonucleotide +188 to +262 (Fig. 3A, lane 2). The specificity of this gel shift complex was demonstrated by competition analysis. Binding was inhibited by the addition of a 30- to 60-fold excess of cold specific competitor (lanes 3 and 4). To determine the nucleotide specificity of binding, a series of oligonucleotides containing various sequences between +188 and +262 were synthesized and assayed for specific competition (lanes 5 to 20). Interestingly, of all the oligonucleotide competitors used, only the sequences from +195 to +240 were able to specifically compete (lanes 15 and 16). Oligonucleotide +202 to +240 would compete only at higher concentrations (data not shown), which indicates that sequences from +195 to +202 were needed for efficient binding. A summary of the gel shift competition results is presented in Fig. 3B.

Purification and characterization of a protein(s) that interacts with DRE 1. To purify the sequence-specific DNA-binding protein(s) that interacted with DRE 1, we subjected Jurkat nuclear extracts to successive rounds of column chromatography on Sephadex G-25, Sephacryl S-200, FPLC Mono-S, and FPLC DNA affinity columns. After each round, the fractions were analyzed by band shift assay. Specific binding was observed with column fractions eluting at approximately 0.2 to 0.25 M KCl from the DNA affinity column. Using the 0.23 M KCl fraction from the DNA affinity column, fraction 16, we performed *in vitro* transcription assays to check for functional activity. The results of these experiments demonstrate that fraction 16 could specifically transactivate the HTLV-I LTR (Fig. 4A, lanes 1 to 4). The concentration of the template DNA, pU3R-CAT, was kept low (200 ng) so that very little transcription was seen in the absence of any transactivator. The transactivation was concentration dependent. When 160 ng of protein was added to the *in vitro* transcription assay, no increase in the level of HTLV-I transcription was observed (lanes 1 and 2). As the concentration of fraction 16 was increased to 480 ng, a significant increase in HTLV-I transcription was observed (lane 3). When the amount of fraction 16 was further increased, 1,120 ng per reaction, no increase in HTLV-I transcription was observed (lane 4). The specificity of fraction 16 was assayed by using the adenovirus major late (AdML) promoter in parallel reactions. No increase in basal

FIG. 2. RT-PCR analysis of RNA following transfection of Jurkat T cells with HTLV-I deletion mutants. (A) The HTLV-I deletion mutant plasmids were electroporated, and RNA was harvested 24 h later. RT-PCR amplification of the RNA was performed as described in Materials and Methods. The PCR products were run on an agarose gel and Southern blotted with CAT-specific probe. (B) Control RT-PCR analysis of actin mRNA. (C and D) Primer extension analysis of the HTLV-I RNA initiation site. RNA from transfected and control cells was incubated with an HTLV-I-specific [γ - 32 P]ATP-end-labeled antisense primer ($\sim 2 \times 10^6$ cpm), 5'-TGTAACGGCGCAGAAC-3', positioned 263 bases downstream of the RNA initiation site. Then 5 μ g of RNA, isolated from transfected or untransfected Jurkat T cells, was added to each reaction. The conditions for the reverse transcription reaction are described in Materials and methods. Products were phenol-chloroform extracted and analyzed on a denaturing sequence gel next to a sequence reaction from either the M13mp18 (C) or the HTLV-I CAT (D) plasmid. Boldface letters denote nucleotides of HTLV-I DNA that are present in A and C sequence ladders.

activity was observed when the DNA affinity column fraction was added to *in vitro* transcription reactions containing the AdML promoter (lanes 5 to 8).

Fraction 16 was made up of four detectable bands by silver stain. These four proteins, p37, p50, p60, and p100, were excised from the gel, renatured and analyzed by band shift assay. Of these four proteins, only purified p37 (Fig. 4C) was able to bind the probe (+188 to +262) (Fig. 4B, lanes 3 to 6). The binding specificity of p37 was analyzed by using the set of oligonucleotides depicted in Fig. 3B as competitors. The p37 band shift could specifically be competed out by the +195 to +240 oligonucleotide (Fig. 4D, lanes 4 and 5; data not shown). It was found that p37 and fraction 16 both have the same specificity in band shift assay.

Finally, we analyzed the effect of protein binding to oligonucleotides containing point mutations (Fig. 5). The selection of this site for point mutagenesis is discussed below. An oligonucleotide was synthesized from +195 to +240, containing five base point mutations at the following positions: +204 (C to A), +205 (T to G), +210 (C to A), +213 (C to A), and +216 (G to A). Band shift analysis was performed to determine whether this oligonucleotide could compete for p37 binding (Fig. 5). The wild-type +195 to +240 oligonucleotide competed for binding (lane 5), but the five-base point mutant was unable to effectively compete (lane 6). Taken together, these results suggest that p37 specifically binds to DRE 1 and may in fact allow higher basal transcription from the HTLV-I promoter.

DISCUSSION

Several reports have indicated that there are two independent types of transcription control elements in the LTR responsible for maximum gene expression. Elements of one type, characterized by regulatory sequences such as TRE-1 or TRE-2, are located upstream from the core promoter in the U3 region and are responsible for transactivation by Tax₁. Another type of element is present in a fragment of 300 bp derived from the R and 5' half of the U5 region. The function of this second element does not require Tax₁. In addition, no extensive homology was found in sequences between the R/U5 fragment and the U3 fragment of the HTLV-I LTR. These structural and functional properties of the second domain suggest that it is a unique control element for gene expression distinct from the Tax₁ responsive element (31, 33). The R regions in the LTRs of HTLV-I and BLV are considerably longer (228 bp) than those of other retroviruses (13 to 100 bp). These longer R regions may be due to the acquisition of sequences to enhance gene expression.

In this study, the effect of sequences at the 3' end of the region that are required for basal transcriptional activity has been analyzed. The sequence that allows maximal detection of the reporter CAT enzyme is from +202 to +246. DRE 1 is functional in several different cell types, including Jurkat T cells and HeLa cells. Therefore, the factor(s) that binds to this sequence and regulates basal transcription of HTLV-I

FIG. 3. Band shift analysis of the protein(s) that binds to the HTLV-I R/U5 DRE 1. (A) An oligonucleotide containing sequences from +188 to +262 relative to the HTLV-I RNA transcription site was used as a probe. The specificity of the gel shift complex was checked by competition with either the same cold competitor (lanes 3 and 4) or other sequences between +188 and +262 (lanes 5 to 20). (B) Summary of the band shift assays presented in panel A.

may be constitutively expressed in most cell types. A specific and stable band shift complex was detected with sequences from +195 to +240. In contrast, sequences from +195 to +235 and +202 to +240 either failed to compete or competed poorly. The results suggest that sequences +195 to +202 and +235 to +240 are important for the specific and stable interaction of cellular proteins with DRE 1. In addition, specific point mutations at bases +204, +205, +210, +213, and +216 inhibited protein binding. Thus, the interaction of p37, and possibly other cellular proteins such as p50, p60, and p100, with this regulatory sequence may be

complex and require extensive regions of the 45-nucleotide regulatory element.

When the DRE 1 DNA-binding factor was purified from Jurkat T cells, we found that a protein of approximately 37 kDa was sufficient to give the proper band shift, which could be specifically competed out by cold competitor (+195 to +240). This protein was excised out of a denaturing polyacrylamide gel, renatured, and used for band shift analysis. It is important to note that although gel-purified p37 could specifically bind DNA, the protein did not increase HTLV-I basal transcription in vitro. We detected increased HTLV-I transcription only from the DNA affinity column fraction 16, which contained four partially purified proteins. The inability of p37 to increase transcription could be due to several factors, including improper folding during renaturation to a protein conformation that is not transcriptionally active. In addition, it is possible that p37 is the DNA-binding subunit of a multisubunit transcription complex and requires proteins such as p50, p60, or p100 to establish a competent transcriptional complex.

Our present analysis demonstrates that the DNA sequence element DRE 1, which was defined through functional in vivo transcription experiments, coincides with a binding site for a cellular transcription factor. Furthermore, addition of the DRE 1 DNA-binding fraction to an in vitro transcription assay results in the specific stimulation of HTLV-I transcription. We cannot, however, rule out the possibility that the regulatory sequence does not contain an overlapping RNA regulatory element. Precedents from HIV argue for caution, since several published reports on the HIV TAR element originally assumed the presence of a DNA element. Experimental procedures to unambiguously delineate RNA regulatory sequences are currently in progress.

The HTLV-I LTR sequence +195 to +240 was scanned against the GenBank data base for homology to other transcriptional regulatory elements. The strongest sequence homology was a 15-base match (100%) from +202 to +217 of the HTLV-I R region to the binding site for human transcription factor TFIID in the 5S rRNA gene H4 (5'-CAAGCAGGGTCAGGC-3'). It is interesting that human TFIID has a molecular mass of approximately 35 kDa and binds to a 51-bp region (+45 to +96) in the transcriptionally active 5S gene (10, 19, 37). TFIID has a number of nucleotide contact sites in the 51-bp region of the 5S gene. Interestingly, base substitution mutation of these contact nucleotides in DRE 1 leads to a loss of p37 protein binding. Two independent lines of investigation, however, suggest that p37 is not the classic TFIID. In vitro gel shift analysis showed that purified *Xenopus* TFIID failed to bind to the DRE 1 oligonucleotide. Although human TFIID would be a more suitable control for these binding studies, cloned human TFIID is not available. In addition, when antiserum to *Xenopus* TFIID was used, no reactivity of the purified p37 was detected in a Western immunoblot analysis. The fact that not all antibodies to *Xenopus* TFIID cross-react with human TFIID, however, complicates the interpretation of the latter experiment. It is possible that the DRE 1-specific factor is one of several proteins within a TFIID-like family, regulating basal promoter activity of the HTLV-I promoter. It is also possible that other transcription factors are required for the functional interaction of TFIID with the HTLV-I LTR. Previously, it has been demonstrated that polymerase II transcription factors, such as TFIID and octamer-binding factor, function in a polymerase III transcription unit (6, 23, 41). It is interesting to speculate that the opposite may also be true, i.e., that classic polymerase III transcription factors

FIG. 4. Analysis of DNA affinity fraction 16. (A) In vitro transcription analysis with fraction 16 and either pU3R (HTLV-I) or AdML promoters. Transcription reactions were set up as described in Materials and Methods. Increasing amounts of the oligonucleotide affinity chromatography fraction (160 ng, lanes 2 and 5; 480 ng, lanes 3 and 6; 1,120 ng, lanes 4 and 8) were added to transcription reactions. Following a 60-min incubation at 30°C, the [32 P]UTP-labeled RNA transcripts were purified and analyzed on a denaturing 4% acrylamide-urea gel. The HTLV-I- and AdML-specific transcripts were 285 and 375 bases, respectively (arrows). (B) Band shift analysis of p37, p50, p60, and p100 isolated from an SDS-polyacrylamide gel. Fraction 16 from the oligonucleotide affinity column was denatured and electrophoresed on an SDS-gel. Subsequently, a portion of the gel was stained; regions of the gel containing p37, p50, p60, and p100 were excised; and proteins were eluted and renatured as described in Materials and Methods. An aliquot of each protein sample was then incubated with the +188 to +262 oligonucleotide probe, and the band shift complexes were analyzed by electrophoresis in a neutral gel. Lane 1, which serves as the positive control, was incubated with fraction 16. (C) The purity of p37 was analyzed by SDS-PAGE. m.w., molecular mass. (D) Specificity of p37 binding to the DRE 1 regulatory sequence. Band shift competition assays were performed to determine the specificity of the interaction between p37 and DRE 1. The DRE 1 probe was incubated alone or in the presence of a 25-fold excess of cold oligonucleotide competitor.

FIG. 5. Mutational analysis of DRE 1. The wild-type probe (+188 to +262) was incubated alone (lane 2) or with a 25-fold excess (300 ng) of a competitor oligonucleotide containing sequences from +188 to +262 (lane 3), +195 to +214 (lane 4), or +195 to +240 (lane 5) or the +195 to +240 five-base mutant containing base substitutions at +204 (C to A), +205 (T to G), +210 (C to A), +213 (C to A), and +216 (G to A) (lane 6).

may be functionally active in a polymerase II transcription unit. This possibility awaits further careful analysis and attention by directly testing human TFIIIA protein and its possible interaction with the polymerase II HTLV-I DRE 1 sequence.

ACKNOWLEDGMENTS

We thank the members of the Laboratory of Molecular Virology for helpful and constructive discussions and Alan Wolffe for generously providing purified *Xenopus* TFIIIA.

ADDENDUM IN PROOF

We have obtained a partial cDNA done for a protein that binds to DRE 1. The clone was obtained from a Jurkat cDNA library by using ³²P-labeled +195 to +240 oligonucleotide. This clone has been sequenced and resembles an inverted CCAAT box binding protein. We are currently pursuing a functional assay that would explore the possibility of this factor's involvement in HTLV-I basal transcription.

REFERENCES

- Ahmed, Y. F., G. M. Gilmartin, S. M. Hanly, J. R. Nevins, and W. C. Greene. 1991. The HTLV-I Rex response element mediates a novel form of mRNA polyadenylation. *Cell* **64**:727-737.
- Bohan, C. A., F. Kashanchi, B. Ensoli, L. Buonaguro, K. Boris-Lawrie, and J. N. Brady. 1992. Analysis of Tat transactivation of human immunodeficiency virus transcription *in vitro*. *Gene Expression* **2**:391-408.
- Bosselut, R., J. F. Duvall, A. Gegonne, M. Bailly, A. Hemar, J. N. Brady, and J. N. Ghysdael. 1990. The product of the c-ets-1 protooncogene and the related Ets2 protein act as transcriptional activators of the long terminal repeat of human T-cell leukemia virus, HTLV-I. *EMBO J.* **9**:3137-3144.
- Bosselut, R., F. Lim, P.-C. Romond, J. Frampton, J. N. Brady, and J. Ghysdael. 1992. Myb protein binds to multiple sites in the human T-cell lymphotropic virus type 1 long terminal repeat and transactivates LTR-mediated expression. *Virology* **186**:764-769.
- Brady, J., K.-T. Jeang, J. Duvall, and G. Khoury. 1987. Identification of p40x-responsive regulatory sequences within the human T-cell leukemia virus type I long terminal repeat. *J. Virol.* **61**:2175-2181.
- Carbon, P., S. Murgo, J.-P. Ebel, A. Krol, G. Tebb, and I. W. Mattaj. 1987. A common octamer motif binding protein is involved in the transcription of U6 snRNA by RNA polymerase II and U2 snRNA by RNA polymerase II. *Cell* **51**:71-79.
- Charnay, P., R. Treisman, P. Mellon, M. Chao, R. Axel, and T. Maniatis. 1984. Differences in human α - and β -globin gene expression in mouse erythroleukemia cells: the role of intragenic sequences. *Cell* **38**:251-263.
- Ciminale, V., G. N. Pavlakis, D. Derse, C. P. Cunningham, and B. K. Felber. 1992. Complex splicing in the human T-cell leukemia virus (HTLV) family of retroviruses: novel mRNAs and proteins produced by HTLV type I. *J. Virol.* **66**:1737-1745.
- Derse, D., and J. W. Casey. 1986. Two elements in the bovine leukemia virus long terminal repeat that regulate gene expression. *Science* **231**:1437-1440.
- Emerson, B. M., and R. G. Roeder. 1984. DNA sequences and transcription factor interactions of active and inactive forms of mammalian 5 S RNA genes. *J. Biol. Chem.* **259**:7926-7935.
- Felber, B. K., H. Paskalis, D. Kleinman-Ewing, F. Wong-Staal, and G. N. Pavlakis. 1985. The pX protein of HTLV-1 is a transcriptional activator of its long terminal repeat. *Science* **229**:675-679.
- Fujii, M., P. Sassone-Corsi, and I. M. Verma. 1988. c-fos promoter trans-activation by the tax1 protein of human T-cell leukemia virus type I. *Proc. Natl. Acad. Sci. USA* **85**:8526-8530.
- Fujisawa, J., M. Toita, T. Yoshimura, and M. Yoshida. 1991. The indirect association of human T-cell leukemia virus tax protein with DNA results in transcriptional activation. *J. Virol.* **65**:4525-4528.
- Gitlin, S. D., R. Bosselut, A. Gegonne, J. Ghysdael, and J. N. Brady. 1991. Sequence-specific interaction of the Ets1 protein with the long terminal repeat of the human T-lymphotropic virus type I. *J. Virol.* **65**:5513-5523.
- Hanly, S. M., L. T. Rimsky, M. H. Malim, J. H. Kim, J. Hauber, M. Duc Dodon, S.-Y. Le, J. V. Maizel, B. R. Cullen, and W. C. Greene. 1989. Comparative analysis of the HTLV-I Rex and HIV-1 Rev trans-regulatory proteins and their RNA response elements. *Genes Dev.* **3**:1534-1544.
- Hultmark, D., R. Klemenz, and W. J. Gehring. 1986. Translational and transcriptional control elements in the untranslated leader of the heat-shock gene *hsp22*. *Cell* **44**:429-438.
- Inoue, J. N., M. Yoshida, and M. Seiki. 1987. Transcriptional (p40^x) and post-transcriptional (p27^{x-111}) regulators are required for the expression and replication of human T-cell leukemia virus type 1 genes. *Proc. Natl. Acad. Sci. USA* **84**:3653-3657.
- Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* **83**:5889-5893.
- Klevit, R. 1991. Recognition of DNA by Cys₂, His₂ zinc fingers. *Science* **253**:1367, 1393.
- Krauskopf, A., O. Resnekov, and Y. Aloni. 1990. A cis downstream element participates in regulation of *in vitro* transcription initiation from the P38 promoter of minute virus of mice. *J. Virol.* **64**:354-360.
- Leung, K., and G. Nabel. 1988. HTLV-1 trans-activator induces interleukin-2 receptor expression through an NF- κ B-like factor. *Nature (London)* **333**:776-778.
- Lilienbaum, A., M. D. Dodon, C. Alexandre, L. Gazzolo, and D. Paulin. 1990. Effect of human T-cell leukemia virus type I Tax protein on activation of the human vimentin gene. *J. Virol.* **64**:256-263.
- Lobo, S. M., J. Lister, M. L. Sullivan, and N. Hernandez. 1991. The cloned RNA polymerase II transcription factor IID selects

- RNA polymerase III to transcribe the human U6 gene in vitro. *Genes Dev.* 5:1477-1489.
24. Manley, J. L., A. Fire, A. Cano, P. A. Sharp, and M. L. Gefter. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci. USA* 77:3855-3859.
 25. Mansour, S. L., T. Grodzicker, and R. Tijan. 1986. Downstream sequences affect transcription initiation from the adenovirus major late promoter. *Mol. Cell. Biol.* 6:2684-2694.
 26. Marriott, S. J., P. F. Lindholm, K. M. Brown, S. D. Gitlin, J. F. Duvall, M. F. Radonovich, and J. N. Brady. 1990. A 36-kilodalton cellular transcription factor mediates an indirect interaction of human T-cell leukemia/lymphoma virus type I Tax₁ with a responsive element in the viral long terminal repeat. *Mol. Cell. Biol.* 10:4192-4201.
 27. Merrill, G. F., R. M. Harland, M. Groudine, and S. McKnight. 1984. Genetic and physical analysis of the chicken *tk* gene. *Mol. Cell. Biol.* 4:1769-1776.
 28. Miyatake, S., M. Seiki, M. Yoshida, and K. I. Arai. 1988. T-cell activation signals and human T-cell leukemia virus type I-encoded p40^{*} protein activate the mouse granulocyte/macrophage colony-stimulating factor gene through a common DNA element. *Mol. Cell. Biol.* 8:5581-5587.
 29. Mondesert, G., C. Tribouley, and C. Keding. 1992. Identification of a novel downstream binding protein implicated in late-phase-specific activation of the adenovirus major late promoter. *Nucleic Acids Res.* 20:3881-3889.
 30. Montagne, J., C. Beraud, I. Crenon, G. Lombard-Platet, L. Gazzolo, A. Sergeant, and P. Jalinot. 1990. Tax1 induction of the HTLV-1 21 bp enhancer requires cooperation between two cellular DNA-binding proteins. *EMBO J.* 9:957-964.
 31. Nakamura, M., K. Ohtani, Y. Hinuma, and K. Sugamora. 1988. Functional mapping of the activity of the R region in the human T-cell leukemia virus type I long terminal repeat to increase gene expression. *Virus Genes* 2:147-155.
 32. Nyborg, J. K., M. H. Matthews, J. Yucel, L. Walls, W. T. Golde, W. S. Dynan, and W. Wachsman. 1990. Interaction of host cell proteins with the human T-cell leukemia virus type-I transcriptional control region. *J. Biol. Chem.* 265:8237-8242.
 33. Ohtani, K., M. Nakamura, S. Saito, T. Noda, Y. Ito, K. Sugamora, and Y. Hinuma. 1987. Identification of two distinct elements in the long terminal repeat of HTLV-1 responsible for maximum gene expression. *EMBO J.* 6:389-393.
 34. Osbourne, T. F., D. N. Arvidson, E. S. Tyau, M. Dunsworth-Browne, and A. J. Berk. 1984. Transcription control region within the protein-coding portion of adenovirus E1A genes. *Mol. Cell. Biol.* 4:1293-1305.
 35. Paskalis, H., B. K. Felber, and G. N. Pavlakakis. 1986. Cis-acting sequences responsible for the transcriptional activation of human T-cell leukemia virus type I constitute a conditional enhancer. *Proc. Natl. Acad. Sci. USA* 83:6558-6562.
 36. Ratner, L. 1990. Regulation of expression of the c-sis proto-oncogene. *Nucleic Acids Res.* 17:4101-4115.
 37. Seifart, K. H., L. Wang, R. Waldschmidt, D. Jahn, and E. Wingender. 1989. Purification of human transcription factor IIIA and its interaction with a chemically synthesized gene encoding human 5 S rRNA. *J. Biol. Chem.* 264:1702-1709.
 38. Seiki, M., A. Hikikoshi, and M. Yoshida. 1990. The U5 sequence is a cis-acting repressive element for genomic RNA expression of human T-cell leukemia virus type I. *Virology* 176:81-86.
 39. Shimotohno, K., M. Takano, T. Teruuchi, and M. Miwa. 1986. Requirement of multiple copies of a 21-nucleotide sequence in the U3 regions of human T-cell leukemia virus type 1 and type 2 long terminal repeats for trans-acting activation of transcription. *Proc. Natl. Acad. Sci. USA* 83:8112-8116.
 40. Siekevitz, M., M. B. Feinberg, N. Holbrook, F. Wong-Staal, and W. C. Greene. 1987. Activation of interleukin 2 and interleukin 2 receptor (Tac) promoter expression by the trans-activator (*tat*) gene product of human T-cell leukemia virus, type I. *Proc. Natl. Acad. Sci. USA* 84:5389-5393.
 41. Simmen, K. A., J. Bernues, H. D. Parry, H. G. Stunnenberg, A. Berkenstam, B. Cavallini, J. M. Egly, and I. W. Mattaj. 1991. TFIID is required for *in vitro* transcription of the human U6 gene by RNA polymerase III. *EMBO J.* 10:1853-1862.
 42. Sodroski, J. G., C. A. Rosen, and W. A. Haseltine. 1984. Trans-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. *Science* 226:177-179.
 43. Stenlund, A., G. L. Bream, and M. R. Botchan. 1987. A promoter with an internal regulatory domain is part of the origin of replication in BPV-1. *Science* 236:1666-1671.
 44. Tendler, C., S. Greenberg, W. Blattner, A. Manns, E. Murphy, T. Fleisher, B. Hanchard, O. Morgan, J. Burton, D. Nelson, and T. Waldmann. 1990. Transactivation of interleukin 2 and its receptor induces immune activation in human T-cell lymphotropic virus type I-associated myelopathy: pathogenic implications and a rationale for immunotherapy. *Proc. Natl. Acad. Sci. USA* 87:5218-5222.
 45. Theill, L. E., O. Wiborg, and J. Vuust. 1987. Cell-specific expression of the human gastrin gene: evidence for a control element located downstream of the TATA box. *Mol. Cell. Biol.* 7:4329-4336.
 46. Watanabe, T., K. Yamaguchi, K. Takatsuki, M. Osame, and M. Yoshida. 1990. Constitutive expression of parathyroid hormone-related protein gene in human T-cell leukemia virus type I (HTLV-I) carriers and adult T-cell leukemia patients that can be trans-activated by HTLV-I tax gene. *J. Exp. Med.* 172:759-765.
 47. Yang, J. Q., E. F. Remmers, and K. B. Marcu. 1986. The first exon of the *c-myc* proto-oncogene contains a novel positive control element. *EMBO J.* 5:3553-3562.