Role of Flanking E Box Motifs in Human Immunodeficiency Virus Type 1 TATA Element Function

S.-H. IGNATIUS OU, LEÓN F. GARCÍA-MARTÍNEZ, EYVIND J. PAULSEN,
AND RICHARD B. GAYNOR*

*Division of Molecular Virology, Departments of Medicine and Microbiology, University of Texas
Southwestern Medical Center at Dallas, Dallas, Texas 75235-8594*

Received 6 May 1994/Accepted 7 August 1994

Human immunodeficiency virus type 1 (HIV-1) gene expression is dependent on a number of *cis*-acting DNA elements present in the HIV-1 long terminal repeat. Previous studies have demonstrated that the TATA element is critical for basal and Tat-induced HIV-1 gene expression. The HIV-1 TATA region has an unusual structure in that the TATA sequence is flanked by two palindromic sequence motifs (CANNTG) known as E boxes which can serve as binding sites for the basic helix-loop-helix (bHLH) class of DNA-binding proteins. In this study, we performed site-directed mutagenesis of both the TATA and the flanking E box sequences of HIV-1. We also substituted the sequences flanking the adenovirus E3 promoter TATA sequence for those flanking the HIV-1 TATA sequence. Constructs were assayed for their levels of basal and Tat-induced gene expression by both *in vitro* transcription and transient expression assays. Both the TATA box and flanking sequences including the E box motifs were found to be important in modulating both basal gene expression and Tat-induced HIV-1 gene expression. Gel retardation analysis demonstrated that binding of both the recombinant TATA-binding protein (TBP) and the TFIID fraction which contains both TBP and TBP-associated factors was dependent primarily on the TATA element. However, competition analysis suggested that the E boxes may play a role in stabilizing the binding of TFIID but not recombinant TBP. Two proteins representing different classes of bHLH proteins, E47 and AP-4, were assayed for their ability to bind to the flanking E box motifs. We isolated a cDNA clone encoding the complete AP-4 protein and demonstrated that both AP-4 and E47 bound specifically to the 3' E box motif, which contains sequences that correspond to the consensus binding site (CAGCTG). Gel retardation analysis indicated that the binding of AP-4 to the E boxes excluded the binding of TBP to the TATA box. These studies are consistent with a model in which different classes of cellular bHLH proteins may be involved in regulating HIV-1 TATA element function by either inhibiting or promoting the assembly of different preinitiation transcriptional complexes.

The regulation of human immunodeficiency virus type 1 (HIV-1) gene expression depends on a number of *cis*-acting elements in the HIV-1 long terminal repeat (LTR) (17). Mutagenesis of several DNA regulatory elements including NF- κ B (44), Sp1 (21, 26), and TATA (1, 16, 27, 35, 46) has indicated their importance in modulating basal and Tat-induced HIV-1 gene expression. In addition to these DNA elements, the structure of the *trans*-activating region (TAR) RNA is also important for high level of gene expression in response to Tat (17). The primary effect of Tat is to stimulate gene expression by increasing the rate of transcriptional elongation, although it is also likely to increase transcriptional initiation (12, 18, 31, 34, 37, 38). Thus, DNA sequences in the HIV-1 promoter such as the TATA element, which is the site of assembly of the preinitiation complex, may help in specifying the binding of particular classes of transcriptional complexes that are uniquely responsive to the effects of Tat (1, 34, 36, 38).

The TATA sequence binds a complex of cellular factors designated TFIID. TFIID is composed of the 38-kDa TATA-binding protein (TBP) which directly binds to the TATA element (19, 48) and at least eight associated cellular factors known as TBP-associated factors, or TAFs (6, 10, 40, 62). Though the functions of most of the TAFs have not been

elucidated, it is thought that they mediate interactions between TBP with either upstream or downstream DNA-binding factors. A functional role for these proteins has been demonstrated in the case of TAF110, which mediates interactions between Sp1 and TBP (23), and TAF250 (56), which modulates TBP activation. Interactions of TBP with the general transcription factors TFIIB and TFIIA are critical in promoting the stepwise assembly of the transcriptional initiation complex (59). Thus, the TFIID complex plays a critical role in determining interactions with both upstream DNA-binding proteins and general transcription factors.

The HIV-1 TATA element is involved in regulating the levels of both basal gene expression and Tat-induced gene expression (1, 16, 27, 35, 46). Both the TATA sequence itself and sequences flanking this motif are important in regulating the degree of *tat* activation. The HIV-1 TATA region is unusual in that it is flanked by two sequence motifs known as E boxes which in other promoters are capable of binding the basic helix-loop-helix (bHLH) class of proteins which include both tissue-specific and constitutive transcription factors (43). The bHLH domain forms two amphipathic α helices, separated by an intervening loop (5, 7, 13, 43). Mutational analysis indicates that the amphipathic helices mediate protein dimerization while basic residues at the amino terminus of the first helix mediate DNA binding (13). The bHLH proteins interact *in vitro* to form stable homodimers or heterodimers with DNA-binding activity (28). The question of whether the DNA sequences in the HIV-1 TATA element can specify the assembly of unique transcription complexes that are responsive to

* Corresponding author. Mailing address: Division of Molecular Virology, Department of Internal Medicine, University of Texas Southwestern Medical School, 5323 Harry Hines Blvd., Dallas, TX 75235-8594. Phone: (214) 648-7570. Fax: (214) 648-8862.

activation by Tat (35) arises. In this study we investigated the potential role of the TATA box and the sequences flanking TATA in the levels of both in vitro gene expression and in vivo gene expression from the HIV-1 LTR. We also characterized the binding properties of either TBP alone or TBP complexed with the TAFs to the HIV-1 TATA box. In addition, we determined the binding of two different classes of bHLH-binding proteins, AP-4 and E47, to the sequences flanking the HIV-1 TATA box. The results of these studies suggest that the E box sequence motifs flanking the HIV-1 TATA element are important in regulating HIV-1 TATA element function.

MATERIALS AND METHODS

Cells and transfections. Jurkat cells were maintained in RPMI with 10% fetal calf serum. HeLa spinner cells were maintained in suspension (S) minimal essential medium containing 5% newborn calf serum. LTR α 3 HeLa cells (62) were also grown in S minimal essential medium but in the presence of 200 μ g of G418 per ml. For transient expression assays, aliquots containing 10^7 Jurkat cells (in 0.5 ml of complete RPMI) were gently mixed with 5 μ g of various HIV-1 LTR templates. In the case of Tat transactivation, 2 μ g of a plasmid containing the Rous sarcoma virus promoter upstream of the *tat* gene was electroporated. The plasmid DNA sample transfections were performed by electroporation using a GIBCO/BRL Cell-Porator at 1,180 μ F and 250 V (field strength, 625 V/cm). The transfected cells were then diluted into 9 ml of complete RPMI. After 36 h of culture, the cells were harvested by centrifugation and the cell pellets were washed twice with 5 ml of phosphate-buffered saline. Jurkat extracts were normalized for protein concentrations. The extracts from transfections containing the *tat* protein were diluted 25-fold compared with extracts lacking Tat, and the chloramphenicol acetyltransferase (CAT) assays were performed as described elsewhere (16). All transfections were performed at least three times, the percentage conversion was quantified on a Phosphorimager, and the results were plotted with the standard deviations shown.

Construction of HIV-1 LTR TATA and E box mutants. Site-directed mutagenesis and PCR were used to generate the various HIV-1 LTR CAT expression vectors used in this study. To generate the HIV-1-simian virus 40 (SV40) TATA LTR CAT plasmid, oligonucleotides 5'-GCAAAAAGCAGCTGC TAAATATGCA-3' and 5'-CTCCTTACGCATCTGTGCGG T-3' were used with PCR to generate a DNA fragment extending from -179 to -18 in the wild-type HIV-1 LTR CAT plasmid. This fragment was cloned into a *Sma*I-cut pUC19 vector, and DNA sequence analysis was performed. An *Eco*RI-*Pvu*II fragment from the latter vector and a *Pvu*II-*Hind*III (-18 to +80) HIV-1 LTR fragment were cloned into an *Eco*RI-*Hind*III-cut CAT expression vector. To generate the HIV-1 NTATA LTR CAT plasmid, oligonucleotides 5'-CAG CTGCTTCGCTGCAGCATCTGAGG-3' and 5'-CTCCTTACGCATCTGTGCGGT-3', were used in PCR. Subsequent cloning steps were the same as those performed for the generation of the HIV-1-SV40 TATA LTR CAT plasmid.

Oligonucleotides 5'-CTGCTTATATGCAGCATCGAAG GCACG-3' and 5'-CTCCTTACGCATCTGTGCGGT-3' were used to generate the HIV-1 LTR CAT 5' E box mutant as described above. To construct the HIV-1 LTR CAT 3' E box mutant plasmid, oligonucleotides 5'-TCGCGACTTATATG CAGCATCTGAGG-3' and 5'-CTCCTTACGCATCTGTGCGGT-3' were used in PCR with the wild-type HIV-1 LTR CAT plasmid as the template. The PCR fragment was cloned into pUC19 cut with *Sma*I and verified by DNA sequence analysis,

and an *Eco*RI-*Nru*I fragment was ligated to a *Pvu*II-*Hind*III fragment from the HIV-1 LTR. To generate the HIV-1 LTR CAT 5' and 3' E box mutants, oligonucleotides 5'-CGACT TATATGCAGCAT-3' and 5'-CTCCTTACGCATCTGTGCGGT-3' were used in PCR with the HIV-1 LTR CAT 5' E box mutant as the template. Cloning of this fragment with the *Pvu*II-*Hind*III fragment and CAT expression vector were performed as for the 3' E box mutant.

Site-directed mutagenesis (Amersham) was used to generate the HIV-1 LTR CAT constructs containing the adenovirus early region 3 (E3) sequences. To construct the 5' flanking sequence, oligonucleotide 5'-GGGACTGGGGAGTGGCGT GCCTCCCGGGCAGGGTATAAGCAGCTGCTTTTTCGCTG-3' was annealed to a single-stranded HIV-1 LTR template and confirmed following in vitro mutagenesis by DNA sequence analysis. An *Eco*RI-*Hind*III fragment was then cloned into an HIV-1 LTR CAT expression vector. The HIV-1 LTR CAT 3' E3 flanking-sequence plasmid was also generated by site-directed mutagenesis using oligonucleotide 5'-AG ATGCTGCATATAACTCACCTGCTTTTTCGCTGT-3'. To construct the HIV-1 LTR CAT plasmid with both 5' and 3' E3 flanking-sequence motifs, the 3' E3 oligonucleotide 5'-CCG GGCAGGGTATAACTCACCTGCTTTTTCGCTGT-3' and the M13 universal reverse primer 5'-AGCGGATAACAATT TCACACAGGA-3' were used in PCR with an HIV-1 LTR CAT plasmid containing 5' E3 flanking sequences as the template. This PCR fragment in pUC19 was restricted with *Sma*I-*Hind*III and ligated with an *Eco*RI-*Sma*I fragment from the HIV-1 LTR CAT 5' E3 flanking plasmid to generate the HIV-1 LTR CAT constructs extending from -179 to +80 with E3 sequences flanking the TATA element.

Isolation of a full-length AP-4 cDNA. Oligonucleotides 5'-GGGGATCCAGGAAAACAGAGAAAGAAGTGAT AG-3' and 5'-GGGCTGGGGGGGTAGTCAGGGAAGC-3' were used to isolate a partial AP-4 cDNA clone by PCR from a Jurkat cDNA library (25). To isolate the full-length AP-4 cDNA, a 378-bp *Sma*I-*Fsp* I cDNA fragment from the partial AP-4 cDNA was random-prime labeled (Pharmacia) and used as a probe to screen a λ EXLox HeLa spinner cell cDNA library. After approximately 10^6 plaques were screened, 10 positive plaques were isolated after tertiary screening, and 3 of the 10 were independent cDNAs, as judged by PCR analysis of the phage DNA. All three clones were cloned into pET, DNA sequence analysis was performed on both strands by dideoxy sequencing methods, and all clones contained the same in-frame stop codon and putative initiating methionine codon with a well-conserved Kozak sequence (GCTATGG).

Gel retardation analysis. Recombinant TBP used in gel retardation assays was purchased from Promega (47). A HeLa nuclear extract, prepared from LTR α 3 HeLa cells (62), which contains an epitope-tagged TBP, was applied to a phosphocellulose column, and fractions designated A to D were eluted with 0.1, 0.3, 0.5, and 1.0 M KCl, respectively (39). The D fraction, which contained an influenza virus epitope-tagged TBP, was detected by Western blot (immunoblot) analysis using the 12CA5 monoclonal antibody with the ECL detection system (Amersham) (62). Both AP-4 (25) and E47 (22, 29) proteins were generated in a rabbit reticulocyte TNT T7 transcription and translation system (Promega). The E47 plasmid, which has been previously described (53), was cloned downstream of the T7 promoter in pGEM 3. AP-4 was cloned in frame as a *Sma*I fragment into the pTM-1 expression vector, which is flanked by T7 promoter and terminator sequences (11). E47 (24) and AP-4 (25) rabbit polyclonal antibodies were generated with glutathione S-transferase fusions of these proteins. For experiments with these antibodies a 1/10 dilution of

immune serum was added to gel retardation assays for 10 min at room temperature after the binding reaction was completed.

Two oligonucleotides, 5'-GCGTGCCTCAGATGCTGCATATAAGCAGCTGCTTTTT-3' and 5'-AAAAAGCAGCTGCTTATATGCAGCATCTGAGGCACGC-3', corresponding to the HIV-1 LTR sequence from -46 to -10 were annealed, labeled with [γ - 32 P]ATP, and used as a probe in gel retardation analysis. DNA-binding reactions for TBP, TFIID, and the phosphocellulose A column competition were carried out in 25 mM KCl-20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid)-KOH (pH 7.9)-0.1 mM EDTA-2 mM spermidine-0.5 mM dithiothreitol-10% glycerol-100 μ g of bovine serum albumin per ml-100 ng of poly(dG-dC) at 30°C for 30 min. For competition experiments, proteins were incubated in the presence of the specific competitor DNA at 30°C for an additional 10 min prior to the addition of the probe. Protein-DNA complexes were resolved by electrophoresis at 125 V at 4°C in a 5% native polyacrylamide gel containing 0.5 $\times$ Tris-borate-EDTA, 1 mM EDTA, and 0.05% Nonidet P-40. The amount of protein used was 1 μ l (TBP) or 3.5 μ g (TFIID or phosphocellulose A fraction). DNA binding for AP-4 and E47 was carried out in 10 mM Tris (pH 7.9)-50 mM KCl-0.1 mM EDTA-0.5 mM dithiothreitol-10% glycerol-(for AP-4 only) 2 μ g of sonicated *Escherichia coli* DNA (Sigma) with 1 μ l of AP-4 or E47 at room temperature for 20 min. Competition was done at room temperature for an additional 10 min before addition of the probe. Protein DNA complexes were resolved by electrophoresis using 6% native polyacrylamide gel and 0.5 $\times$ Tris-borate-EDTA at a constant voltage of 125 V at 4°C. The sequences of the coding strand of the competitor oligonucleotides that were used included 5'-GCGTGCCTCAGATGCTGCAGCGAAGCAGCTGCTTTTT-3' (HIV-1 NTATA), 5'-GCGTGCCTTCGATGCTGCATATAAGTCGCTGCTTTTT-3' (HIV-1 double E box mutant), 5'-GCGTGCCTTCGATGCTGCAGCGAAGTCGCTGCTTTTT-3' (HIV-1 NTATA double E box mutant), 5'-GCGTGCCTTCGATGCTGCATATAAGCAGCTGCTTTTT-3' (HIV-1 5' E box mutant), and 5'-GCGTGCCTCAGATGCTGCATATAAGTCGCTGCTTTTT-3' (HIV-1 3' E box mutant).

In vitro transcription. In vitro transcription reactions were performed as described elsewhere (58) with a HeLa cell nuclear extract (8) and 0.2 μ g of both an *Nco*I-linearized HIV-1 LTR CAT template, which generated a 620-bp runoff transcript, and an *Spe*I-cut HIV-1 LTR template containing a substitution of the TAR loop sequences between +31 and +34, which generated an 1,100-bp runoff transcript (16). Reactions in 45 μ l were incubated for 60 min at 30°C in 10 mM HEPES (pH 7.9)-50 mM KCl-0.1 mM EDTA-0.25 mM dithiothreitol-5 mM MgCl₂-0.06 mM each ATP, CTP, and UTP-0.016 mM GTP-10 μ Ci of [α - 32 P]GTP (3,000 Ci/mM)-10% glycerol. The reactions were stopped by the addition of 400 μ l of a stop solution containing 7.0 M urea, 0.1 M Tris HCl (pH 7.4), 0.5 M sodium chloride, 0.5% sodium dodecyl sulfate, 10 mM EDTA, and 25 μ g of yeast tRNA per ml. The samples were extracted with phenol-chloroform and precipitated with ethanol. The transcription products were analyzed on a 4% polyacrylamide gel (7 M urea, 1 $\times$ Tris-borate-EDTA) run at 60 mW for 5 h. The dried gel was exposed to a Molecular Dynamics Phosphorimager screen, and the reaction products were quantitated by using Imagequant software from Molecular Dynamics.

For in vitro transcription with recombinant Tat produced in bacteria, Tat with a glutathione *S*-transferase moiety attached to the 3' portion was purified by glutathione agarose chromatography (51), as was a *tat* mutant with mutations in three critical cysteine residues which has previously been shown to

	5' E Box	TATA Box	3' E Box
HIV WT	CT CAGATG CTGCA	TATAA	G CAGCTG CTTT
HIV SV40 TATA	CT CAGATG CTGCA	TATTTA	G CAGCTG CTTT
HIV NTATA	CT CAGATG CTGCA	gcgAA	G CAGCTG CTTT
HIV TATA 5' HLH (-)	CT tcGATG CTGCA	TATAA	G CAGCTG CTTT
HIV TATA 3' HLH (-)	CT CAGATG CTGCA	TATAA	G tcGCTG CTTT
HIV TATA 5'/3' HLH (-)	CT tcGATG CTGCA	TATAA	G tcGCTG CTTT
HIV 5'/3' E3 TATA	CT CCCGGG CAGGG	TATAA	G CACCTG CTTT
HIV 5'/3' E3 NTATA	CT CCCGGG CAGGG	gcgAA	G CACCTG CTTT

FIG. 1. Sequences of the wild-type (WT) and mutant HIV-1 TATA elements. The sequences flanking the HIV-1 TATA element extending from -40 to -12 relative to the transcription start site are shown, and the TATA sequence and E box motifs are indicated. Mutants which were analyzed include those containing mutations of the TATA sequence (SV40 TATA and NTATA [a nonsense sequence]) the 5' E box [5' HLH (-)], the 3' box [3' HLH (-)], and both the 5' and the 3' boxes [5'/3' HLH (-)] and those containing substitutions of the HIV-1 TATA element with the flanking sequences of the adenovirus early region 3 TATA element (5'/3' E3 TATA and 5'/3' E3 NTATA).

be very defective in transcriptional activation (15). Cotransfection assays using the *tat*-glutathione *S*-transferase construct cloned into a eucaryotic expression vector indicated that it activated HIV-1 LTR gene expression to a level similar to that for an expression vector containing *tat* alone. The strategy for observing Tat transactivation was that described by Marciniak et al. (37). RNA was extracted and subjected to electrophoresis on a 6% denaturing polyacrylamide gel as described previously, and autoradiography was performed.

RESULTS

Architecture of the HIV-1 TATA element. The HIV-1 TATA element contains the sequence TATAA, extending from nucleotide -27 to -23 relative to the transcription initiation site. Flanking the HIV-1 TATA sequence are two direct repeats between -38 and -33 (CAGATG) and -21 and -16 (CAGCTG) which are homologous with the consensus E box sequence CANNTG (43). The E box motif serves as the binding site for both the bHLH (43) and the bHLH leucine zipper (bHLH-ZIP) (4, 20, 55) groups of DNA-binding proteins (28, 42). Among the bHLH proteins are transcription factors such as c-Myc, MyoD, myogenin, and E12/E47, while the bHLH-ZIP group includes USF, TEF-3, and AP-4 (28). The exact DNA sequence to which different bHLH and bHLH-ZIP proteins bind optimally is determined primarily by the central dinucleotide present within the CANNTG motif (3, 52). For example, c-Myc binds to CACGTG (2), whereas AP-4 binds to CAGCTG (25).

Because of the diversity of proteins that can potentially bind to the E box motifs, we first determined the role of the TATA sequence and flanking E box motifs on the regulation of HIV-1 gene expression by performing mutagenesis of each of these elements. As shown in Fig. 1, TATA region mutants which changed the HIV-1 TATA to sequences found in the SV40 TATA element or to a non-TATA sequence were constructed. The second set of mutants involved disrupting either the 5' E box motif, the 3' E box motif, or both the 5' and the 3' E boxes. The third set of mutants involved replacing the HIV-1 TATA region with the same region from another viral promoter, the adenovirus early region 3 promoter, a procedure which in essence substituted the sequences flanking TATA but preserved the TATA motif (33). The adenovirus early region 3

FIG. 2. Gene expression of HIV-1 TATA element mutations. HIV-1 LTR CAT constructs extending from -179 to +80 in the HIV-1 LTR fused to the CAT gene and containing TATA element mutations shown in Fig. 1 were electroporated into Jurkat cells in both the presence and the absence of Tat. CAT assays were performed at 36 h postelectroporation, and the fold Tat induction for each construct was calculated by determining the percent of CAT conversion in both the presence and the absence of Tat. The values shown were calculated from the values obtained in three different transfections. Constructs tested included the wild-type (WT) HIV-1 LTR and those containing a substitution of the HIV-1 TATA sequence with the SV40 TATA sequence or a nonsense TATA sequence (NTATA) a mutation of either the 5' E box [5' HLH (-)], the 3' E box [3' HLH (-)], or both the 5' and the 3' E boxes [5'/3' HLH (-)], or a substitution of the HIV-1 TATA element with the adenovirus early region 3 TATA and flanking sequences (5'/3' E3 TATA) or with a mutated TATA sequence (5'/3' E3 NTATA).

promoter contains the same TATA sequence as that in the HIV-1 LTR but contains a different E box sequence (CAC-CTG) located 3' to the TATA sequence and a GC-rich sequence located 5' to the TATA box (33). By analyzing the effects of each of these mutations on both basal gene expression and Tat-induced gene expression and on the binding of both TBP and different bHLH proteins, we could better ascertain the role the HIV-1 TATA box and flanking sequences play in regulating HIV-1 gene expression.

The TATA sequence and the flanking E boxes are important in regulating HIV-1 gene expression in Jurkat cells. Constructs containing either mutated TATA or mutated flanking E box sequences were introduced into HIV-1 LTR constructs extending from -179 to +80 and fused to the CAT gene (16). These constructs were introduced into Jurkat cells by electroporation in both the presence and the absence of Tat. CAT assays were performed at 36 h postelectroporation, and the percent conversion of [14 C]chloramphenicol was determined. Fold activation by Tat was determined by comparing the percentage of CAT conversion in both the presence and the absence of Tat. The results of three separate experiments were incorporated into Fig. 2. Mutation of the HIV-1 TATA sequence to the SV40 TATA sequence decreased Tat activation four- to five-

fold, while mutagenizing the TATA sequence to a nonsense TATA sequence essentially eliminated Tat activation (Fig. 2).

Next, we determined whether changes in the E box motifs flanking the HIV-1 TATA sequence influenced the levels of basal and Tat-induced expression. Mutation of the 5' E box had only slight effects on Tat induced gene expression (Fig. 2). However, disruption of either the 3' E box or both the 5' and the 3' E boxes reduced Tat activation about fourfold. Substitution of the HIV-1 TATA element with the TATA box and flanking sequences of the adenovirus E3 promoter also resulted in approximately a fourfold decrease in Tat activation. Incorporating a nonsense TATA sequence into the latter construct nearly eliminated Tat activation (Fig. 2). These results demonstrated the importance of both the TATA and flanking E box sequences in regulating basal and Tat-induced gene expression from the HIV-1 LTR.

The TATA sequence and flanking motifs influence basal and Tat-induced in vitro transcription from the HIV-1 LTR. Since transfection experiments demonstrated that both the TATA and flanking sequences were important in regulating basal and Tat-induced gene expression, we further investigated the role of these regions in in vitro transcription from the HIV-1 LTR. Each of these templates was linearized with *Nco*I and then incubated with a HeLa nuclear extract and ribonucleotides to generate a 620-bp runoff transcript (58). An HIV-1 template containing a mutation of the TAR RNA loop sequences between +31 and +34 was cut with *Spe*I to generate an 1,100-bp runoff transcript and was included in each reaction as a control. This construct has been demonstrated to exhibit marked decreases in activation in response to Tat (16, 37). The results of a typical experiment to analyze the basal levels of gene expression are shown in Fig. 3A. Experiments with each set of mutants were repeated four times, and the results were quantitated by Phosphorimager scanning.

Substitution of the HIV-1 TATA sequence with that of SV40 reduced basal promoter activity approximately fivefold (Fig. 3A, lane 2). Substitution of the HIV-1 TATA sequence with a nonsense sequence GCGAA reduced promoter activity approximately 10-fold (Fig. 3A, lane 3). Constructs which altered the E box motifs flanking the HIV-1 TATA element were also tested. Disruption of the 5' E box motif was found to have only modest effects on the level of in vitro transcription from the HIV-1 LTR (Fig. 3A, lane 4). Disruption of the 3' E box motif resulted in approximately a threefold decrease in in vitro transcription from the HIV-1 LTR (Fig. 3A, lane 5). Constructs with combined mutations of both the 5' and the 3' E box motifs resulted in a level of transcription that was fourfold lower than that seen with the wild-type construct (Fig. 3A, lane 6). Substitution of the adenovirus E3 flanking region for the HIV-1 flanking regions resulted in a fivefold decrease in transcription (Fig. 3A, lane 7), while coupling a TATA mutation with the adenovirus E3 flanking motifs resulted in a 20-fold decrease in the level of transcription from the HIV-1 LTR (Fig. 3A, lane 8).

Next, we used in vitro transcription assays with each of these different HIV-1 templates in the presence of bacterially produced wild-type Tat or a *tat* mutant containing changes in three critical cysteine residues which have been demonstrated to prevent its ability to activate HIV-1 gene expression (15). As shown in Fig. 3B (lanes 1 and 2), Tat activated in vitro transcription from the wild-type HIV-1 LTR construct approximately sevenfold. Changing the HIV-1 TATA sequence to that of SV40 or to a nonsense TATA sequence virtually eliminated both basal gene expression and Tat-induced gene expression (Fig. 3B, lanes 3 to 6). Mutation of the 5' E box did not alter the degree of Tat activation, while mutation of the 3'

FIG. 3. Basal and Tat-induced in vitro transcription of the HIV-1 LTR. (A) In vitro transcription was performed by using a HeLa nuclear extract with an HIV-1 LTR CAT template which was linearized at a unique *Nco*I site to generate a 620-bp runoff transcript. An HIV-1 template containing a wild-type HIV-1 TATA element (lane 1); the SV40 TATA element (lane 2); a nonsense TATA element (lane 3); a mutation of the 5' E box (lane 4), 3' E box (lane 5), or both the 5' and the 3' E boxes (lane 6); the adenovirus early region 3 TATA element (lane 7); or the same element containing a mutated TATA element (lane 8) instead of the wild-type HIV-1 TATA element were used in in vitro transcription assays. A second template, which contained substitutions of the TAR RNA loop sequences between +31 and +34 was cut with *Spe*I to generate an 1,100-bp runoff transcript. (B) Either bacterially produced Tat (odd lanes) or Tat containing a mutation in three critical cysteine residues (even lanes) was added to HIV-1 templates containing a wild-type TATA element (lanes 1 and 2), the SV40 TATA element (lanes 3 and 4), NTATA (a nonsense sequence) (lanes 5 and 6), a mutation of the 5' E box (lanes 7 and 8), a mutation of the 3' E box (lanes 9 and 10), a mutation of both the 5' and the 3' E boxes (lanes 11 and 12), or a substitution of the HIV-1 TATA element with that of adenovirus early region 3 TATA element (lanes 13 and 14) or the adenovirus early region 3 promoter TATA element containing the NTATA sequence (lanes 15 and 16). The HIV-1 TAR RNA loop mutant +31/+34 cut with *Spe*I to generate an 1,100-bp runoff transcript was included in each reaction.

E box reduced Tat activation twofold. Mutation of both the 5' and the 3' E boxes decreased Tat activation fourfold (Fig. 3B, lanes 7 to 12). Substitution of the HIV-1 flanking sequences with those of the adenovirus E3 promoter in either the presence or the absence of a wild-type TATA sequence resulted in constructs with no detectable activation by Tat (Fig. 3B, lanes 13 to 16). Thus, the effects of TATA region mutations on Tat activation were more pronounced in in vitro transcription assays than in transfection assays. These results indicated that both the TATA sequence itself and the flanking sequences are critical in the process of Tat activation of the HIV-1 LTR.

Mutation of the HIV-1 TATA sequence but not of the flanking E box motifs alters TBP binding. The effects of mutations of the TATA sequence and flanking E boxes on HIV-1 gene expression could be explained potentially by decreased binding of TBP to the TATA element. This would be consistent with the observation that a higher binding affinity of TBP to the TATA element correlates with increased promoter activity in in vitro transcription assays (57). Recombinant TBP produced in bacteria was used in gel retardation analysis with labeled oligonucleotides corresponding to sequences extending from -46 to -10 in the HIV-1 LTR (47).

FIG. 4. Gel retardation analysis of TBP binding to wild-type (WT) and mutated HIV-1 TATA elements. Oligonucleotides extending from -45 to -10 in the HIV-1 LTR were used in gel retardation analysis with bacterially produced TBP. Gel retardation was performed with the probe alone (lane 0) or TBP either without competition (lane 1) or with a 10-, 20-, or 50-fold molar excess of unlabeled oligonucleotides corresponding to either wild-type HIV-1 LTR sequences (lanes 2 to 4), mutations of both the 5' and the 3' E boxes (lanes 5 to 7), mutation of the TATA sequence (lanes 8 to 10), or mutations of the 5' and 3' E boxes in addition to the TATA sequence (lanes 11 to 13).

As shown in Fig. 4, TBP bound to these oligonucleotides in gel retardation analysis. Addition of increasing amounts of unlabeled wild-type HIV-1 TATA element oligonucleotides resulted in the competition of the gel-retarded TBP species (Fig. 4, lanes 2 to 4). Oligonucleotides with mutations of both E box motifs also competed effectively for TBP binding (Fig. 4, lanes 5 to 7). However, oligonucleotides containing mutations of the TATA element alone (Fig. 4, lanes 8 to 10) or the TATA element and the flanking E box motifs (Fig. 4, lanes 11 to 13) did not result in significant competition for TBP. These results indicated that mutation of the HIV-1 TATA sequence alone but not of the flanking E box motifs altered the binding of TBP.

TFIID binds specifically to the HIV-1 TATA element. It was important to address whether the E box motifs flanking the TATA sequence might modulate the binding of TBP and TAFs. We used a HeLa cell line, designated LTR α 3, which was stably transformed with a TBP cDNA tagged with the influenza virus hemagglutinin epitope (62). In this way, we could easily assay for the presence of this epitope-tagged TBP using the 12CA5 monoclonal antibody, which is directed against a 12-amino-acid portion of the influenza virus hemagglutinin (14). A HeLa nuclear extract was fractionated on a phosphocellulose column, and the TFIID fraction was eluted with 1.0 M KCl (45). Western blot analysis using the 12CA5 antibody was performed, and the result indicated the presence of the 38-kDa TBP in the 1.0 M KCl eluate from the phosphocellulose column (Fig. 5A, lane 6).

Gel retardation analysis indicated that two gel-retarded species in the TFIID fraction bound specifically to the HIV-1 TATA probe (Fig. 5B). The binding of the faster-migrating species was inhibited by the addition of 12CA5 antibody (data not shown). Both the binding of the faster-migrating gel-retarded species and the binding of the one migrating more slowly were specifically blocked by wild-type HIV-1 oligonucleotides (Fig. 5B, lanes 2 to 4) but not by oligonucleotides containing mutations in the HIV-1 TATA element (Fig. 5B,

FIG. 5. Binding of TFIID to the HIV-1 TATA element. (A) A nuclear extract prepared from a HeLa cell line, LTR α 3, which express an epitope-tagged TBP gene was eluted from a phosphocellulose column with 1.0 M KCl and analyzed by Western blot with the 12CA5 monoclonal antibody. Nuclear extract alone (lane 1); phosphocellulose column flowthrough (lane 2); and 0.1, 0.3, 0.5, and 1.0 M KCl (lanes 3 to 6, respectively) eluates from the phosphocellulose column were analyzed by Western blot analysis with the 12CA5 monoclonal antibody. Lane M, molecular weight markers. (B) Oligonucleotides corresponding to sequences in the HIV-1 LTR which extend from -45 to -10 (lane 0) were used in gel retardation analysis with 3.5 μ g of the 1.0 M KCl eluate from the phosphocellulose column (lane 1). Competition analysis was performed with either a 5-, a 20-, or a 50-fold molar excess of unlabeled oligonucleotides corresponding to the HIV-1 wild-type TATA element (lanes 2 to 4), mutations of both the 5' and the 3' E boxes (lanes 5 to 7), mutation of the TATA box (lanes 8 to 10), or mutations of both TATA and the 5' and 3' E boxes (lanes 11 to 13). Arrows, gel-retarded complexes that bind to the TATA motif.

lanes 8 to 13). Oligonucleotides containing mutations in the flanking E box motifs which maintained the TATA sequence competed similarly to the wild-type HIV-1 TATA element oligonucleotides for binding of both retarded species (Fig. 5B, lanes 5 to 7). Oligonucleotides containing mutations of the TATA sequences alone competed poorly for both retarded species (Fig. 5B, lanes 8 to 10), while mutations of both the TATA and E boxes consistently gave less competition than oligonucleotides containing mutations of the TATA element alone (Fig. 5B, lanes 11 to 13). These results indicated that the TATA sequence was critical for binding of both gel-retarded species in the TFIID fraction, though the flanking E boxes may contribute to the affinity of this binding.

E47 binds preferentially to the 3' E box in the HIV-1 TATA element. It was also important to determine whether different bHLH proteins could bind to either the 5' or the 3' E box motif flanking the TATA sequence in the HIV-1 LTR. If this is the case, it would suggest the possibility that bHLH (29, 43) proteins can either prevent or enhance the binding of the

TFIID complex to the HIV-1 TATA element. Several studies have demonstrated that the primary sequence of the E box influences the binding of different types of bHLH proteins (3, 52). E47 is a ubiquitous bHLH protein, whereas AP-4 is a ubiquitous bHLH-ZIP protein (25), and both bind optimally to the sequence CAGCTG. Therefore, we used one representative protein from these two different classes of bHLH proteins to determine their ability to bind to the flanking E boxes and any potential differences in their binding specificities.

First, we tested whether in vitro-translated E47 protein could bind to the E boxes flanking the HIV-1 TATA sequence. Labeled oligonucleotides corresponding to the wild-type HIV-1 TATA element or these same sequences containing mutations in either the 5' E box, the 3' E box, both the 5' and the 3' E boxes, or the 5' and 3' E boxes and the TATA element were used in gel retardation analysis with either an unprogrammed reticulocyte lysate (Fig. 6A, lanes 6 to 10) or a lysate programmed with E47 RNA (Fig. 6A, lanes 11 to 15). The E47 protein bound well to oligonucleotides corresponding to the HIV-1 wild-type TATA element (Fig. 6A, lane 11) and mutations in either the TATA box (Fig. 6A, lane 12) or the 5' E box (Fig. 6A, lane 13). However, E47 bound poorly to HIV-1 TATA element probes which contained mutations in either the 3' E box (Fig. 6A, lane 14) or both the 5' and 3' E boxes (Fig. 6A, lane 15).

Since several gel-retarded species were seen with the reticulocyte lysate alone and did not markedly change with the different probes (Fig. 6A, lanes 6 to 10), antibodies to either E47 or preimmune sera were added to the gel retardation reaction mixtures to confirm the specificity of E47 binding. As shown in Fig. 6B, by using oligonucleotides corresponding to the HIV-1 TATA element, the presence of E47 antiserum (Fig. 6B, lane 5) but not preimmune serum (Fig. 6B, lane 4) resulted in a supershift of the E47 complex bound to the HIV-1 TATA element. These results indicated that the observed gel-retarded species contained E47 and that E47 bound preferentially to the 3' E box in the HIV-1 TATA element.

AP-4 binds specifically to the 3' E box in the HIV-1 TATA element. The full-length protein coding sequence of AP-4 has not previously been determined (25). We isolated a 378-bp fragment from the published AP-4 cDNA sequence and used this fragment to screen a λ EXLox 10 HeLa cDNA library. Positive plaques were purified and subjected to DNA sequence analysis. A 2.1-kb cDNA whose DNA sequence agreed with the previously published AP-4 sequence except for an additional 17 amino acids with an in-frame initiating methionine and an upstream stop codon (25) was identified. The sequence of the full-length AP-4-coding sequence is shown in Fig. 7. The AP-4 cDNA was then inserted downstream of the T7 promoter in the pTM-1 vector (11) to produce in vitro-translated AP-4 protein in a rabbit reticulocyte lysate.

We tested whether in vitro-translated AP-4 could bind to the HIV-1 TATA region. Gel retardation analysis demonstrated that AP-4 bound to oligonucleotides corresponding to the HIV-1 TATA element (Fig. 8A, lane 3). This binding was inhibited by both wild-type HIV-1 TATA element oligonucleotides (Fig. 8A, lanes 4 to 6) and oligonucleotides with a mutated TATA motif but with the E box sequences intact (Fig. 8A, lanes 7 to 9). However, oligonucleotides containing mutations of both flanking E boxes (Fig. 8A, lanes 10 to 12) or mutations of the TATA sequences in addition to the E boxes (Fig. 8A, lanes 13 to 15) did not compete for the binding of AP-4.

To determine whether AP-4 could bind efficiently to each of the flanking E box motifs, gel retardation and competition analysis was again performed with labeled oligonucleotides

FIG. 6. Gel retardation analysis of E47 binding to wild-type and mutated HIV-1 TATA elements. Gel retardation was performed with probes alone (lanes 1 to 5), 1 μ l of unprogrammed reticulocyte (Retic) lysate (lanes 6 to 10), or 1 μ l of E47 protein translated in a rabbit reticulocyte lysate (lanes 11 to 15). Oligonucleotide probes that were used in gel retardation analysis include those corresponding to the wild-type HIV-1 TATA element (lanes 1, 6, and 11), NTATA (a nonsense sequence) (lanes 2, 7, and 12), a 5' E box mutation (lanes 3, 8, and 13), a 3' E box mutation (lanes 4, 9, and 14), and 5' and 3' E box mutations (lanes 5, 10, and 15). (B) Gel retardation analysis was performed with a wild-type HIV-1 TATA element oligonucleotide (lane 1), 1 μ l of an unprogrammed reticulocyte lysate (lane 2), 1 μ l of E47 protein produced in a reticulocyte lysate (lane 3), or 1 μ l of E47 protein in the presence of either 1 μ l of preimmune rabbit serum (lane 4) or 1 μ l of rabbit polyclonal antibody (Ab) directed against E47 (lane 5).

corresponding to the wild-type TATA element. Oligonucleotides that either corresponded to the wild-type HIV-1 TATA element (Fig. 8B, lanes 4 to 6) or contained mutations in the 5' E box motif competed for AP-4 binding (Fig. 8B, lanes 7 to 9). However, oligonucleotides containing mutations of the 3' E box or both the 5' and the 3' E box motifs failed to compete for AP-4 binding (Fig. 8B, lanes 10 to 15). These results demonstrated that AP-4 bound specifically to the HIV-1 TATA element 3' E box, which contains the consensus binding site for AP-4 (CAGCTG).

Role of AP-4 in TBP binding to the HIV-1 TATA element. Since AP-4 bound preferentially to the 3' E box, the question of whether both TBP and AP-4 could simultaneously bind to the HIV-1 TATA element arose. As shown in Fig. 9A, both bacterially produced TBP (lane 1) and AP-4 produced in a rabbit reticulocyte lysate (lane 3) bound to the HIV-1 TATA element. Increasing the quantity of AP-4 added to the gel retardation assay resulted in competition for TBP binding to the HIV-1 TATA element (Fig. 9A, lanes 5 to 9). The decrease in TBP binding was not due to the degradation of TBP, since the maximum amount of the unprogrammed lysate used did not affect binding of TBP (Fig. 9A, lane 4). To demonstrate that this inhibition requires the binding of both TBP and AP-4 to the HIV-1 TATA element, gel retardation was performed with the HIV-1 TATA element oligonucleotides containing mutations in the flanking E box motifs. As shown in Fig. 9B, increasing the concentration of AP-4 did not induce competition for the binding of TBP to HIV-1 TATA element oligonucleotides containing mutations in the E box sequences (Fig. 9B, lanes 5 to 7). These results demonstrated that AP-4 was able to compete with TBP for binding to the HIV-1 TATA element. Furthermore, the addition of AP-4 protein was found to inhibit *in vitro* transcription from the HIV-1 LTR (data not shown), indicating that high concentrations of bHLH proteins

can compete with TFIID for binding to the HIV-1 TATA element.

Cellular factors bind to the 3' HLH motif in a nuclear extract. To determine if cellular factors present in a HeLa nuclear extract were capable of binding to the E box motifs flanking the HIV-1 TATA sequence, oligonucleotides corresponding to this region were used in gel retardation analysis with different fractions of a HeLa nuclear extract chromatographed on a phosphocellulose column. Five major gel-retarded species were observed with the 0.1 M KCl eluate of a HeLa nuclear extract fractionated on a phosphocellulose column by using oligonucleotides corresponding to the wild-type HIV-1 TATA element (Fig. 10A, lane 3). Four of these gel-retarded species were also seen with oligonucleotides containing a mutated TATA element (Fig. 10A, lane 4). A probe which mutated the 5' E box did not markedly alter the positions or intensities of these gel-retarded species (Fig. 10B, lane 2). However, mutation of either the 3' E box motif alone or both the 5' and the 3' E box motifs resulted in the loss of one major gel-retarded species (Fig. 10B, lanes 3 and 4). These results were consistent with the fact that cellular factors present in the nuclear extract were capable of binding specifically to the 3' E box motif in the HIV-1 TATA element. Samples of a HeLa nuclear extract fractionated on a phosphocellulose column and eluted with 0.3, 0.5, or 1.0 M KCl did not give appreciable differences in the pattern of gel-retarded species with all the labeled probes used (data not shown).

Finally, we wanted to determine whether AP-4, which, we had shown, bound specifically to the 3' E box, had a mobility in gel retardation analysis similar to that of the nuclear protein which bound to the 3' E box. As shown in Fig. 10C, AP-4 synthesized in a rabbit reticulocyte lysate had a mobility similar to that of the cellular factor found in the nuclear extract that

GCATCTGGGACCGCCGATTCTGCACCTCCGTCCGGCGTCCCTTTGATTCGGATTTCATCTTCATTCTCCGGCTGATCGCGGACCTGGCTCGTCGACAGGAGGGGGCCGATGCTATGAGGATTTTCATGG 274
M E Y F M

TGCCCATCTCAGAAGGTGCCCTCTTTGCAACATTTTCAGGAAAAAGATGATAGGAGGGCTCTGTAGCTTGCACACATCCACTAACCCCCGAGACTCAGCGGACAGGAGCGCGGATTCGGCGGGAG 411
V P T Q K V P S L Q H F R K T E K E V I G G L C S L A N I P L T P E T Q R D Q E R R I R R E

ATCGCCAACAGCAACGAGCGGAGACGATCGAGAGCATCAACGCGGATTCAGTCCCTCAAGACCCCTCATCCCCACACAGAGCGAGAGAAGCTCAGCAAGGCGAGCCATTCTCCAGCAGACAGCCGAGTACATCTT 548
I A N S H E R R H O S I N A S F O S L K T L I P H T D G E K L S K A A I L G Q T A E Y I F

TCTCTCGGAGCAGGAGACACAGGCTCTTGACAGCAACACAGCTCAAGCGCTTCATCCAGGAGCTGAGCGGCTGCTCCCCAAGCGACGGCGGCGAGGACGAGCAAGGATAGGCTCCCCGGGACCT 685
S L E Q E K T R L L Q Q N T Q L K R F I Q E L S G S P K R R R A E D K D E G C A T C I

GGGAGGACGCAAGGCGGAGGACCTCGCGCGGAGATGATTGAGCTGCGGAGCAGCTGGACAAGGAGCGCTCGGTGCGCATGATGCTGGAGGAGCAGGTGCGCTCGCTGAGGAGCCCATGTACCCGGAAAAGCTC 822
V E D E K A E D L R R E M I E L R Q Q L D K E R S V R M H L E E Q V R S L E A H M Y P E K L

AAGGTGATTGCGCAGCAGGTGCGACTGACGACGAGCAGGAAACAGGTGAGGCTGCTGACACGAGAGAAGCTGGAGCGGAAACAGCAGCAGTGGCGGACCGAGTCTGCCCCCTCGGGCCCCACCCACCCAC 959
K V I A Q Q V Q L Q Q Q Q E Q V R L L H Q E K L E R E Q Q Q L R T Q L L P P P A P T H H P T

GGTGATCTGTGCGCAGCACCCTCTCTCTCTCCACCATCAATGTCGTACCATGGGCGCTCTCGGTCACTGTTTCCACATCCCGCAAAATCTGGACACCATCGGCGAGGCAATCCAGCAGCATCG 1098
V I V P A P P P P P S H I N V I V T M G P S S V I N S V T S R Q N L D T I V Q A I Q H I

AGGGCACCAGGAAAAGCAGGAGCTGAGGAGGAGCAGCGGCGAGCTGTCTGCTGAAGCTGTGCGCAGCTGCCGAGGGCCCCACCTGTGACACCGCTTCGAGCTCGAGGCTCAGACAGTGAGCTGAGGAC 1233
E G T Q E K Q E L E E E Q R R A V I V K P V R S C P E A P T S D T A S D S E A S D S D A H D

CAGAGCGGGAGGAGCGCTCGGGGACGGGAGCTTCCCTGACTACCCCCAGCCCTCTCTCTCTTGGGGCTGAGGGGAGCGGGGAGCCACAGGGAGAGACATGGGCGAATGAGTGAAGAAATTTTACAA 1370
Q S R E E P S G D G E L P

AATTACGATGCTATTGGGCTCTTTTATGACCTCTTTTCAATACTGTAATGACCTTTGAACGAAGCCACTCAACCCGAGGTCCCAGGGCTGGGGTGTGCGAGAGCTGTGGAGCATCGGACCCCGAGGGCGGG 1507

GCTCTGGCCCCGGGGCTGGAGGAAGCTGACACGGAGATGCTGGCTCTCTGTGCAAAAAGCATTTTTCCTTTAAATATGTTTTTAAAGACAGGAAAATTAACAAAACCCAGGTATTCTCTCCCTGCC 1844

AGAGCCAGCTGGGATTGTGACGCTTAATCCCCCTTCTCTCTCTTTTGGGTTTTCTTCTTCTCTTTAAGCACTTACATGGTTGGGGTAAGCATAGGCTGGGGCATCTGGGGGCGGAGGCTCTCCGTTGC 1781

TTCTTGGTGGGGTTTGTGCTGCTGTGCCCCCTTCCCTTCCCATCTCGGCACTAGAATTGCGCACTCTCCACCCCGAGCCCCACCTCTGCTCCAGGCTCATCTTCCACCCCAAAATGTCTGCTCTC 1918

TCTTTTGTGTTTTGTGTTGGTTTTTATTTCTTTTTGGTTTTGCTTCTGTTTTGTTTTGTTTTTCTTTTTTCTTTTTTTTTTTTACAATTTGAGGCTTCTGCTGTTCAAGGAGAAGCTATTATA 2055

TTTTGTAAGAAAGTGGGAGAAAAAAACCAAGAGCCACCGTGCTTTGTAAAGAAACAAATAAGTTTGACTTTGTTTTTAAAAAAA 2149

FIG. 7. Isolation of a cDNA clone encoding the full-length AP-4 protein. A HeLa cDNA library in λ EXLox was screened with a 378-bp end-labeled probe from a portion of the partial AP-4 coding sequence. Positive phages identified by this screening were subject to DNA sequence analysis, and both the 2,149-bp nucleotide sequence and the amino acid sequence for the full-length AP-4 protein are shown. The two helices in the AP-4 protein are shaded.

bound to the 3' E box (Fig. 10C, lanes 3 and 4). Rabbit polyclonal antibody directed against the AP-4 protein, but not preimmune serum, was able to prevent the binding of this species present in a HeLa nuclear extract (Fig. 10D, lanes 4 to 6). There was no effect of this AP-4 antibody on either of the two more slowly migrating gel-retarded species (Fig. 10D, lane 6). Thus, the species present in the HeLa nuclear extract which bound to the 3' E box in the HIV-1 TATA element is likely to be immunologically related to AP-4.

DISCUSSION

Both viral and cellular promoters transcribed by RNA polymerase II may frequently differ in the structure of their TATA elements, though it is likely that they all bind TBP (59). For instance, recombinant TBP has been shown to bind to the SV40 late promoter, dihydrofolate reductase, and terminal deoxynucleotidyltransferase promoters which do not contain consensus TATA sequences (57). However, differences in the TATA sequence can alter the ability of promoter elements to respond to diverse transcriptional activation signals. The heat shock promoter (hsp70) is strongly activated by the adenovirus E1A protein when its native TATA sequence is present, but it is not activated when an SV40 early promoter TATA sequence is substituted (50). Both the native hsp70 and the SV40 TATA sequences retain inducibility by heat shock, indicating that differences in the TATA element may be important in determining the responsiveness of a promoter to different transcriptional activation signals (50). Similar experiments have demonstrated that the structure of the TATA sequence was important for responsiveness to a muscle-specific enhancer (54). Thus, the nature of the TATA sequence is critical for

mediating responses to different transactivators and upstream enhancers. A study of sequence motifs flanking TATA elements indicates that they may also play a role in regulating gene expression. For instance, in the human papillomavirus virus type 18 p105 promoter, it has been demonstrated that the bovine papillomavirus E2 protein is able to bind to sequence motifs overlapping the TATA element. The binding of E2 to these sequences prevents the subsequent binding of TBP and inhibits expression from this promoter (8a). Therefore, there is precedent for both the TATA sequence itself and the sequences surrounding TATA altering the level of promoter-directed gene expression.

Our studies using both transient transfection assays and *in vitro* transcription assays demonstrated that both the HIV-1 TATA sequence itself and sequences flanking TATA influenced the level of both basal transcription and Tat-induced transcription. Changing the HIV-1 TATA sequence to that of SV40 markedly decreased both basal gene expression and Tat-induced gene expression. Changing the HIV-1 TATA sequence to a nonsense TATA sequence nearly abolished both basal gene expression and Tat-induced gene expression. These results indicated that the primary sequence of the TATA motif was important in regulating gene expression. Since the affinity of TBP binding to nonconsensus TATA sequences correlates with its ability to activate gene expression (57), the decreased gene expression seen with the SV40 TATA sequence may be due to altered binding of the TFIID complex (TBP and TAFs) or to the assembly of a different TFIID complex. Mutation of the HIV-1 TATA sequence to a nonsense TATA sequence severely diminished the binding of both TBP and TFIID. Other studies which investigated the role of the HIV-1 TATA element in both basal gene expression and Tat-induced gene

FIG. 8. Gel retardation analysis of AP-4 binding to wild-type and mutated HIV-1 TATA elements. Oligonucleotides spanning the HIV-1 TATA element and extending from -45 to -10 were used in gel retardation analysis with an in vitro-transcribed and -translated AP-4. Gel retardation is shown for the probe alone (lane 1), $1 \mu\text{l}$ of an unprogrammed reticulocyte lysate (lane 2), and $1 \mu\text{l}$ of a reticulocyte lysate programmed with AP-4 RNA alone (lane 3) or with a 100-, 200-, or 500-fold molar excess of unlabeled oligonucleotides corresponding to the wild-type HIV-1 TATA element (lanes 4 to 6), a mutation of the TATA sequence (lanes 7 to 9), a mutation of both the 5' and the 3' E boxes (lanes 10 to 12), or a mutation of both the 5' and the 3' E boxes in addition to the TATA sequence (lanes 13 to 15). (B) Competition experiments similar to those of panel A except that competition was performed with oligonucleotides corresponding to the wild-type HIV-1 TATA element (lanes 4 to 6), a mutated 5' E box (lanes 7 to 9), a mutated 3' E box (lanes 10 to 12), or mutated 5' and 3' E boxes (lanes 13 to 15).

expression have also indicated that the structure of the HIV-1 TATA element is critical for optimal Tat activation (1, 35, 47).

The HIV-1 promoter is a rare example of an important RNA polymerase II promoter in which the TATA box is flanked in close proximity by binding sites for sequence-specific transcription factors. Thus, we investigated the role of E box sequences flanking the HIV-1 TATA in *in vivo* and *in vitro*

gene expression. Site-directed mutations of both the 5' and the 3' E boxes resulted in marked decreases in basal and Tat-induced gene expression, indicating a role for these elements in HIV-1 TATA element function. Constructs maintaining the HIV-1 TATA sequence but containing substitutions of both E boxes and flanking sequences with those from the adenovirus E3 TATA element resulted in defects in gene expression that

FIG. 9. AP-4 competes with TBP for binding to the HIV-1 TATA element. (A) Gel retardation analysis was performed with oligonucleotides corresponding to a wild-type HIV-1 probe by using in vitro-translated AP-4 and bacterially produced TBP. The probe alone (lane 0), $1 \mu\text{l}$ of TBP alone (lane 1), an unprogrammed reticulocyte lysate (lane 2), $1 \mu\text{l}$ of AP-4 alone (lane 3), $1 \mu\text{l}$ of TBP with $5 \mu\text{l}$ of an unprogrammed reticulocyte lysate (lane 4), and $1 \mu\text{l}$ of TBP with increasing amounts (1 to 5 μl) of AP-4 (lanes 5 to 9) were used in gel retardation analysis. (B) Gel retardation was performed with HIV-1 TATA element oligonucleotides containing mutations in both the 5' and the 3' E boxes. Lane 0, probe alone; lane 1, $1 \mu\text{l}$ of TBP alone; lane 2, $1 \mu\text{l}$ of an unprogrammed reticulocyte lysate; lane 3, $1 \mu\text{l}$ of AP-4 alone; lane 4, $1 \mu\text{l}$ of TBP with $5 \mu\text{l}$ of an unprogrammed lysate; lanes 5 to 7, $1 \mu\text{l}$ of TBP with increasing amounts of (1, 3, and 5 μl , respectively). AP-4.

FIG. 10. Gel retardation analysis of the binding of nuclear proteins to the HIV-1 TATA element. (A) Gel retardation analysis of oligonucleotides corresponding to the wild-type HIV-1 TATA sequence or a mutated TATA sequence in the absence (lanes 1 and 2) or the presence (lanes 3 and 4) of a HeLa nuclear extract eluted with 0.1M KCl from a phosphocellulose column containing 3.5 μ g of protein. (B) A HeLa nuclear extract was used in gel retardation analysis with oligonucleotide probes corresponding to the wild-type HIV-1 TATA element (lane 1), a mutation of the 5' E box (lane 2), a mutation of the 3' E box (lane 3), or mutations of both the 5' and the 3' E boxes (lane 4). (C) Oligonucleotides corresponding to the wild-type HIV-1 TATA element were used in gel retardation analysis in the absence of an extract (lane 1), in the presence of 1 μ l of an unprogrammed reticulocyte lysate (lane 2), in the presence of a reticulocyte lysate containing 1 μ l of AP-4 (lane 3), or in the presence of a HeLa nuclear extract (lane 4). (D) Oligonucleotides corresponding to the HIV-1 TATA element were used in gel retardation assays in the absence of an extract (lane 1), in the presence of a 1/10 dilution of 1 μ l of either preimmune rabbit serum (lane 2) or AP-4 rabbit polyclonal antibody (lane 3), or in the presence of a HeLa nuclear extract alone (lane 4) or together with either preimmune serum (lane 5) or AP-4 antibody (lane 6). About 3.5 μ g of the HeLa nuclear extract was used for each binding reaction. Arrows, gel-retarded species which comigrate with AP-4.

were even more marked than those seen with mutation of the E box sequences alone. Other transcriptional regulatory elements, in addition to the TATA sequence and E boxes, are also likely to be involved in regulating basal and Tat activation. For instance, the HIV-1 LTR contains a bipartite initiator element located between -5 and $+9$ and $+27$ and $+40$ (9, 60), and mutations of these elements decrease gene expression. Both USF and a related factor designated TFII-I bind to the HIV-1 initiator and regulate its function (49). In addition, the transcription factor UBP-1/LBP-1, which binds to TAR DNA, can also bind to low-affinity sites adjacent to the HIV-1 TATA element and repress gene expression (27, 32). Thus, the proteins which bind to the sequences flanking the HIV-1 TATA box are an important determinant for subsequent activation by Tat.

The question of how the E box sequences flanking the HIV-1 TATA element might influence basal and Tat-induced gene expression arises. Gel retardation analysis using templates mutated in the flanking E boxes showed that the ability of purified bacterially produced TBP to bind to the HIV-1 TATA element was not altered. However, it was noted that oligonucleotides containing mutations of both the HIV-1 TATA sequence and E boxes competed less well for binding the TFIID complex than oligonucleotides containing a mutation of the TATA element alone. This result, which was seen reproducibly in a number of different experiments, suggested that the flanking E box motifs may help to regulate the binding affinity of the TBP-TAF complex while not affecting the binding of TBP alone. Thus, it is possible that the binding of the TBP-TAF complex may be regulated by both the sequence of the TATA box and the sequences flanking TATA. Different TBP-TAF complexes may exist, and a subset of these may have a preferential ability to bind to the HIV-1 TATA element and respond directly by binding of the *tat* protein or perhaps by more indirect mechanisms (30, 35).

A previous study demonstrated the binding of a variety of bHLH proteins to three potential E box motifs in the HIV-1 LTR between -303 and -298 , -166 and -161 , and -21 and -16 (61). However, the functional role of these elements and their binding specificity were not analyzed. We investigated the binding properties of both bHLH (E47) and bHLH-ZIP (AP-4) transcriptional factors with respect to flanking E box motifs in the HIV-1 TATA region. Mutagenesis results indicated that the 3' E box, compared with the 5' box, is more important in determining the level of both basal gene expression and Tat-induced gene expression. The 3' E box motif in the HIV-1 TATA region contains the palindromic CAGCTG sequence. Studies have indicated that both bHLH and bHLH-ZIP proteins have higher affinities for a perfect palindromic E box sequence (3). E47 was used in this study because it is a ubiquitous bHLH protein that has a high specificity for the sequence CAGCTG. Similarly, AP-4 is a ubiquitous bHLH-ZIP protein and the only known bHLH-ZIP protein that binds preferentially to the consensus CAGCTG (25, 41). Other bHLH-ZIP proteins such as c-Myc, USF, and TEF-3 all bind to the consensus sequence CACGTG (28). Our gel retardation analysis results indicated that both E47 and AP-4 bound preferentially to the 3' E box motif. Furthermore, we found that a cellular protein(s) present in a HeLa nuclear extract with an antigenicity similar to that of AP-4 also bound preferentially to the 3' E box, though we were unable to detect E47 binding to this sequence. These results suggested that the 3' E box flanking the HIV-1 TATA box can serve as a consensus high-affinity binding site for bHLH proteins with resultant modulation in the level of HIV-1 gene expression.

In vitro transcription and transfection studies suggested that factors binding to the E box motifs flanking the HIV-1 TATA box should result in increases in gene expression. However, we observed that the binding of AP-4 excluded the binding of TBP to the HIV-1 TATA box. In vitro transcription analysis with

the HIV-1 LTR indicated that the addition of purified AP-4 protein repressed gene expression in the presence of an intact 3' E box. Thus, a variety of constitutive bHLH proteins may potentially act as repressors of HIV-1 expression. In Jurkat cells which were used in our transfection studies, the TFIID complex was likely present in excess such that competition by AP-4 did not occur. However, it is possible that in other cell types either AP-4 or other bHLH proteins may function to repress HIV-1 gene expression by preventing the binding of the TFIID complex. The increase in gene expression mediated by the E box motifs suggests a model in which a specific group of as yet undefined bHLH-ZIP proteins, which are likely associated with the complex of TBP and TAFs, binds to the E boxes flanking the HIV-1 TATA box to promote the assembly of the TFIID complex. Thus, both positive factors and negative factors may be capable of binding to the E boxes flanking the HIV-1 TATA element. Regardless of the mechanisms of transcription modulation by these proteins, we have clearly established the importance of the flanking E box motifs in regulating the levels of basal and Tat-induced HIV-1 gene expression.

ACKNOWLEDGMENTS

We thank Ty Lawrence and Brian Finley for the preparation of the manuscript, Brian Finley for the artwork, Arnie Berk for the gift of LTR α 3 cells, Hai-Ling Hsu and Richard Baer for E47 antibody, Foon Wu for HeLa nuclear extract, and Joe Garcia for encouragement and suggestions.

This work was funded by grants from the NIH, the Veterans Administration, and the Robert Welch Foundation. S.-H. I. Ou is supported by the UT Southwestern Medical Scientist Training Program grant GM08014 and the Perot Family Foundation.

REFERENCES

- Berkhout, B., and K.-T. Jeang. 1992. Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* **66**:139–149.
- Blackwell, T. K., L. Kretzner, E. M. Blackwood, R. N. Eisenman, and H. Weintraub. 1990. Sequence-specific DNA binding by the c-Myc protein. *Science* **250**:1149–1151.
- Blackwell, T. K., and H. Weintraub. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**:1104–1110.
- Blackwood, E. M., and R. N. Eisenman. 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* **251**:1211–1217.
- Brennan, T. J., T. Chakraborty, and E. N. Olson. 1991. Mutagenesis of the myogenin basic region identifies an ancient protein motif critical for activation of myogenesis. *Proc. Natl. Acad. Sci. USA* **88**:5675–5679.
- Chiang, C. M., H. Ge, Z. Wang, A. Hoffmann, and R. G. Roeder. 1993. Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *EMBO J.* **12**:2749–2762.
- Davis, R. L., P. F. Cheng, A. B. Lassar, and H. Weintraub. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**:733–746.
- Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475–1489.
- Dostatni, N., P. F. Lambert, R. Sousa, J. Ham, P. M. Howley, and M. Yaniv. 1991. The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex. *Genes Dev.* **5**:1657–1671.
- Du, H., A. L. Roy, and R. G. Roeder. 1993. Human transcription factor USF stimulates transcription through the initiator elements of the HIV-1 and the Ad-ML promoters. *EMBO J.* **12**:501–511.
- Dynlacht, B. D., T. Hoey, and R. Tjian. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**:563–576.
- Elroy, S. O., T. R. Fuerst, and B. Moss. 1989. Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. *Proc. Natl. Acad. Sci. USA* **86**:6126–6130.
- Feinberg, M. B., D. Baltimore, and A. D. Frankel. 1991. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. *Proc. Natl. Acad. Sci. USA* **88**:4045–4049.
- Ferre, D. A., G. C. Prendergast, E. B. Ziff, and S. K. Burley. 1993. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature (London)* **363**:38–45.
- Field, J., J. Nikawa, D. Broek, B. MacDonald, L. Rodgers, I. A. Wilson, R. A. Lerner, and M. Wigler. 1988. Purification of a RAS-responsive adenyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. *Mol. Cell. Biol.* **8**:2159–2165.
- Garcia, J. A., D. Harrich, L. Pearson, R. Mitsuyasu, and R. B. Gaynor. 1988. Functional domains required for tat-induced transcriptional activation of the HIV-1 long terminal repeat. *EMBO J.* **7**:3143–3147.
- Garcia, J. A., D. Harrich, E. Soultanakis, F. Wu, R. Mitsuyasu, and R. B. Gaynor. 1989. Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. *EMBO J.* **8**:765–778.
- Gaynor, R. 1992. Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS* **6**:347–363.
- Graeble, M. A., M. J. Churcher, A. D. Lowe, M. J. Gait, and J. Karn. 1993. Human immunodeficiency virus type-1 trans-activator protein, tat, stimulates transcriptional read-through of distal terminator sequences *in vitro*. *Proc. Natl. Acad. Sci. USA* **90**:6184–6188.
- Greenblatt, J. 1991. Roles of TFIID in transcriptional initiation by RNA polymerase II. *Cell* **66**:1067–1070.
- Gregor, P. D., M. Sawadogo, and R. G. Roeder. 1990. The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes Dev.* **4**:1730–1740.
- Harrich, D., J. Garcia, F. Wu, R. Mitsuyasu, J. Gonzalez, and R. Gaynor. 1989. Role of SP1-binding domains in *in vivo* transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* **63**:2585–2591.
- Henthorn, P., M. Kiledjian, and T. Kadesch. 1990. Two distinct transcription factors that bind the immunoglobulin enhancer microE5/kappa 2 motif. *Science* **247**:467–470.
- Hoey, T., R. O. Weinzierl, G. Gill, J. L. Chen, B. D. Dynlacht, and R. Tjian. 1993. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* **72**:247–260.
- Hsu, H. L., I. Wadman, and R. Baer. 1994. Formation of *in vivo* complexes between the TAL1 and E2A polypeptides of leukemic T cells. *Proc. Natl. Acad. Sci. USA* **91**:3181–3185.
- Hu, Y. F., B. Luscher, A. Admon, N. Mermod, and R. Tjian. 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. *Genes Dev.* **4**:1741–1752.
- Jones, K. A., J. T. Kadonaga, P. A. Luciw, and R. Tjian. 1986. Activation of the AIDS retrovirus promoter by the cellular transcription factor, SP1. *Science* **232**:755–759.
- 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.
- Kadesch, T. 1993. Consequences of heteromeric interactions among helix-loop-helix proteins. *Cell Growth Differ.* **4**:49–55.
- Kamps, M. P., C. Murre, X. H. Sun, and D. Baltimore. 1990. A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. *Cell* **60**:547–555.
- Kanshanchi, F., G. Piras, M. F. Radonovich, J. F. Duvall, A. Fattaey, C. M. Chiang, R. G. Roeder, and J. N. Brady. 1994. Direct interaction of human TFIID with the HIV-1 transactivator tat.

- Nature (London) **367**:295–299.
31. Kao, S.-Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin. 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by *tat* gene product. *Nature* (London) **330**:489–493.
 32. Kato, H., M. Horikoshi, and R. G. Roeder. 1991. Repression of HIV-1 transcription by a cellular protein. *Science* **251**:1476–1479.
 33. Kornuc, M., S. Kliewer, J. Garcia, D. Harrich, C. Li, and R. Gaynor. 1990. Adenovirus early region 3 promoter regulation by E1A/E1B is independent of alterations in DNA binding and gene activation of CREB/ATF and AP1. *J. Virol.* **64**:2004–2013.
 34. Laspia, M., A. Rice, and M. B. Mathews. 1989. HIV-1 *tat* protein increases transcriptional initiation and stabilizes elongation. *Cell* **59**:283–292.
 35. Lu, X., T. M. Welsh, and B. M. Peterlin. 1993. The human immunodeficiency virus type 1 long terminal repeat specifies two different transcription complexes, only one of which is regulated by Tat. *J. Virol.* **67**:1752–1760.
 36. Lu, Y., M. Stenzel, J. G. Sodroski, and W. A. Haseltine. 1989. Effects of long terminal repeat mutations on human immunodeficiency virus type 1 replication. *J. Virol.* **63**:4115–4119.
 37. Marciniak, R. A., B. J. Calnan, A. D. Frankel, and P. A. Sharp. 1990. HIV-1 Tat protein trans-activates transcription in vitro. *Cell* **63**:791–802.
 38. Marciniak, R. A., and P. A. Sharp. 1991. HIV-1 Tat protein promotes formation of more-processive elongation complexes. *EMBO J.* **10**:4189–4196.
 39. Matsui, T., J. Segall, P. A. Weil, and R. G. Roeder. 1980. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J. Biol. Chem.* **255**:11992–11996.
 40. Meisterernst, M., and R. Roeder. 1991. Family of proteins that interact with TFIID and regulate promoter activity. *Cell* **67**:557–567.
 41. Mermod, N., T. J. Williams, and R. Tjian. 1988. Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcription in vitro. *Nature* (London) **332**:557–561.
 42. Murre, C., P. S. McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**:777–783.
 43. Murre, C., P. S. McCaw, H. Vassin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore. 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**:537–544.
 44. Nabel, G., and D. Baltimore. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* (London) **326**:711–713.
 45. Nakajima, N., M. Horikoshi, and R. G. Roeder. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol. Cell. Biol.* **8**:4028–4040.
 46. Olsen, H. S., and C. A. Rosen. 1992. Contribution of the TATA motif to Tat-mediated transcriptional activation of human immunodeficiency virus gene expression. *J. Virol.* **66**:5594–5597.
 47. Peterson, M. G., N. Tanese, B. F. Bugh, and R. Tjian. 1990. Functional domains and upstream activation properties of cloned human TATA binding protein. *Science* **248**:1625–1630.
 48. Rigby, P. W. 1993. Three in one and one in three: it all depends on TBP. *Cell* **72**:7–10.
 49. Roy, A. L., M. Meisterernst, P. Pognonec, and R. G. Roeder. 1991. Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. *Nature* (London) **354**:245–248.
 50. Simon, M. C., T. M. Fisch, B. J. Benecke, J. R. Nevins, and N. Heintz. 1988. Definition of multiple, functionally distinct TATA elements, one of which is a target in the hsp70 promoter for E1A regulation. *Cell* **52**:723–729.
 51. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**:31–40.
 52. Sun, X. H., and D. Baltimore. 1991. An inhibitory domain of E12 transcription factor prevents DNA binding in E12 homodimers but not in E12 heterodimers. *Cell* **64**:459–470.
 53. Sun, X.-H., N. G. Copeland, N. A. Jenkins, and D. Baltimore. 1991. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol. Cell. Biol.* **11**:5603–5611.
 54. Wefald, F., B. Devlin, and R. S. Williams. 1990. Functional heterogeneity of mammalian TATA-box sequences revealed by interaction with a cell-specific enhancer. *Nature* (London) **344**:260–262.
 55. Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, and S. Hollenberg. 1991. The myoD gene family: nodal point during specification of the muscle cell lineage. *Science* **251**:761–766.
 56. Weinzierl, R. O., B. D. Dynlacht, and R. Tjian. 1993. Largest subunit of *Drosophila* transcription factor IID directs assembly of a complex containing TBP and a coactivator. *Nature* (London) **362**:511–517.
 57. Wiley, S. R., R. J. Kraus, and J. E. Mertz. 1992. Functional binding of the “TATA” box binding component of transcription factor TFIID to the –30 region of TATA-less promoters. *Proc. Natl. Acad. Sci. USA* **89**:5814–5818.
 58. Wu, F., J. Garcia, D. Sigman, and R. Gaynor. 1991. Tat regulates binding of the human immunodeficiency virus trans-activation region RNA loop-binding protein TRP-185. *Genes Dev.* **5**:2128–2140.
 59. Zawel, L., and D. Reinberg. 1992. Advances in RNA polymerase II transcription. *Curr. Opin. Cell Biol.* **4**:488–495.
 60. Zenzie-Gregory, B., and S. T. Smale. 1993. The HIV-1 core promoter lacks a simple initiator element but contains a bipartite activator at the transcription start site. *J. Biol. Chem.* **268**:15823–15832.
 61. Zhang, Y., K. Doyle, and M. Bina. 1992. Interactions of HTF4 with E-box motifs in the long terminal repeat of human immunodeficiency virus type 1. *J. Virol.* **66**:5631–5634.
 62. Zhou, Q., P. M. Lieberman, T. G. Boyer, and A. J. Berk. 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev.* **6**:1964–1974.