The EMBO Journal vol.11 no.7 pp.2573-2581, 1992

Transcription factor AP-2 activates gene expression of HTLV-I

Christian Muchardt, Jacob-S.Seeler, Ajay Nirula, Shiaoching Gong and Richard Gaynor

Division of Molecular Virology, Department of Internal Medicine, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas, USA

Communicated by P.Boulanger

The HTLV-I LTR contains three conserved regulatory elements known as 21 base pair repeats which are required for stimulation of gene expression by the transactivator protein tax. Mutagenesis indicates that the 21 bp repeats can be subdivided into three motifs, A, B and C, each of which influences the level of tax activation. The A site in the 21 bp repeat has strong homology with previously described binding sites for the transcription factor AP-2. We demonstrated that AP-2 mRNA was present in T-lymphocytes and that cellular factors from both non-transformed and transformed T-lymphocytes specifically bound to the consensus motif for AP-2 in each 21 bp. To determine the role of AP-2 in the regulation of the HTLV-I LTR gene expression, we used an AP-2 cDNA in DNA binding and transient expression assays. Gel retardation and methylation interference studies revealed that bacterially produced AP-2 bound specifically and with high affinity to all three 21 bp repeats, and that it required the core sequence AGGC for specific binding. Binding of AP-2 prevented the subsequent binding of members of the CREB/ATF family to an adjacent regulatory motif in the 21 bp repeat. Transfection of an AP-2 expression construct into T-lymphocytes activated gene expression from the HTLV-I LTR. At least two 21 bp repeats were required for high levels of AP-2 activation and mutagenesis of the AP-2 consensus binding sequences in the 21 bp repeats eliminate this activation. These results demonstrate that AP-2 is capable of binding to and activating gene expression from the HTLV-I LTR, and thus may play a role in the pathogenesis of this important human retrovirus.

Key words: 21 bp repeat/AP-2/HTLV-I/*tax*

Introduction

The human T-cell leukemia virus type I (HTLV-I) is the causative agent of adult T-cell leukemia (Poiesz *et al.*, 1980, 1981; Yoshida *et al.*, 1982). HTLV-I is also associated with a progressive demyelinating syndrome known as tropical spastic paraparesis (Gessain *et al.*, 1985; Osame *et al.*, 1986). HTLV-I has been demonstrated to infect $CD4^+$ T-lymphocytes and to immortalize these cells. A 40 kDa protein encoded by HTLV-I known as tax is involved in the transforming properties of this virus (Hinrichs *et al.*, 1987; Wano *et al.*, 1988; Pozzati *et al.*, 1990; Tanaka *et al.*, 1990). In addition, tax is critical for the regulation of viral gene expression (Sodroski *et al.*, 1984; Chen *et al.*, 1985; Felber *et al.*, 1985; Fujisawa *et al.*, 1985; Smith and Greene, 1990). The properties of tax which mediate both its transcriptional and transforming activities have been the subject of intense investigation.

At least two different elements in the HTLV-I LTR are critical for activation of gene expression. One element with the repeated sequence CC(A/T)CC serves as the binding site for the ets family of transcription factors (Bosselut *et al.*, 1990; Gitlin *et al.*, 1991). Both ets-1 and ets-2 have been shown to be capable of activating gene expression of the HTLV-I LTR CAT constructs in cotransfection experiments (Bosselut *et al.*, 1990). The second element which is critical for gene expression of the HTLV LTR is the 21 bp repeat (Fujisawa *et al.*, 1986; Paskalis *et al.*, 1986; Shimotohno *et al.*, 1986; Brady *et al.*, 1987; Rosen *et al.*, 1987). Three of these highly conserved repeats are present in different portions of the HTLV-I LTR and mutagenesis studies indicate that at least two 21 bp repeats are required for a high level of induction by tax (Fujisawa *et al.*, 1986; Brady *et al.*, 1987). Thus, the 21 bp repeat and potentially other elements in the HTLV-I LTR are critical for tax activation.

Mutagenesis of the 21 bp repeat indicates that it can be subdivided into three distinct motifs known as A, B and C (Fujisawa *et al.*, 1989; Montagne *et al.*, 1990). The B motif contains the core sequence TGAGG which serves as a potential binding site for members of the ATF/CREB family (Hai *et al.*, 1989; Yoshimura *et al.*, 1990). The A and C motifs which flank the B motif contain the core recognition sequences AGGC and CCCC respectively. Mutagenesis of the B motif eliminates tax activation (Fujisawa *et al.*, 1989; Montagne *et al.*, 1990), while mutagenesis of both the A and C motifs also result in severe decreases in tax-mediated activation of the HTLV-I LTR (Montagne *et al.*, 1990). Thus, the 21 bp repeats are composed of different motifs each of which may be required for complete activation by the tax protein.

Since tax does not directly bind to the HTLV-I LTR, its role in increasing gene expression is thought to be mediated by interaction with cellular factors bound to the LTR (Marriot *et al.*, 1990; Béraud *et al.*, 1991). A number of different cellular factors can bind to the 21 bp repeats (Altman *et al.*, 1988; Jeang *et al.*, 1988; Nyborg *et al.*, 1988, 1990; Tan *et al.*, 1989a; Montagne *et al.*, 1990; Nyborg and Dynan, 1990; Yoshimura *et al.*, 1990; Béraud *et al.*, 1991). At least eight members of the CREB/ATF family can bind to the core sequence TGACG found in the B motif in the HTLV-I LTR (Hai *et al.*, 1989; Tsujimoto *et al.*, 1990; Yoshimura *et al.*, 1990). Fractionation of cellular extracts and DNase I footprinting reveals that AP-2 or related factors bind over a portion of the 21 bp repeat, though which nucleotides are critical for this binding remain

unclear (Nyborg and Dynan, 1990). In addition, other less well defined factors have also been demonstrated to bind to different motifs in the 21 bp repeat (Tan *et al.*, 1989; Montagne *et al.*, 1990; Nyborg and Dynan, 1990; Yoshimura *et al.*, 1990; Béraud *et al.*, 1991). In this study we explore the binding and regulation of AP-2 (Williams *et al.*, 1988; Williams and Tjian, 1991a and b) to the HTLV-I LTR. We demonstrate that AP-2 mRNA is expressed in T-lymphocytes and that proteins from both T-lymphocytes and HeLa cells bind specifically and with high affinity to the AGGC sequence within the A motif of the HTLV-I LTR 21 bp repeat. Likewise, proteins produced in bacteria from an AP-2 cDNA also binds to this motif and the binding of AP-2 to the A motif in the 21 bp repeat excludes the subsequent binding of members of the CREB/ATF family. Finally, we demonstrate that AP-2 can activate gene expression from the HTLV-I LTR. These studies describe the role of the cellular transcription factor AP-2 in the regulation of HTLV-I LTR gene expression.

Results

HTLV-I LTR 21 bp repeats contain sequences with homology to AP-2 binding sites

A schematic diagram of the position of the three 21 bp repeats in the HTLV-I LTR relative to the transcriptional start site and the nucleotide sequence of each of the three 21 bp repeats are shown in Figure 1A. Binding assays and transfection experiments with mutagenized 21 bp repeat constructs were consistent with the fact that the 21 bp repeat can be further subdivided into three regulatory motifs known as A, B and C, each of which have distinct regulatory properties (Figure 1A) (Fujisawa *et al.*, 1989; Montagne *et al.*, 1990).

Examination of the DNA sequences in the A motif of the 21 bp repeats revealed strong homology to the binding site for the cellular transcription factor AP-2 which has previously been defined in the SV40 and metallothionein enhancers (Figure 1B) (Mitchell *et al.*, 1987; Imagawa *et al.*, 1987). The highest degree of homology with SV40 was found in 21 bp repeat III, with a lower degree of homology found in repeats I and II. The sequence AGGC was completely conserved in the A motif of all three 21 bp repeats.

AP-2 is present in T-lymphocytes

Since HTLV-I can infect and transform T-lymphocytes, we first determined whether AP-2 mRNA was expressed in these cells. Primers were synthesized to regions of the published AP-2 cDNA sequence (Williams *et al.*, 1988) and used in polymerase chain reaction with total RNA isolated from the Jurkat T-lymphocyte cell line. As shown in Figure 2, four products were synthesized in the PCR reaction. These products were all cloned and analyzed by DNA sequencing. The 560 bp product had a sequence identical to the published AP-2 cDNA sequence (Williams *et al.*, 1988). A smaller product of ~290 bp (II) appeared to be an alternatively spliced product of AP-2 mRNA which introduced a stop codon prematurely into the AP-2 coding sequence (Figure 2). Two other products of 170 and 90 bp could not be related to the AP-2 cDNA sequence and were probably PCR artifacts (Figure 2). A similar pattern of PCR products was seen in HUT78 cells (data not shown).

Next we determined whether factors present in either Jurkat cells or an HTLV-I transformed T-lymphocyte cell line, SLB (Koeffler *et al.*, 1984), bound to an AP-2 binding site. Labeled oligonucleotides corresponding to a high affinity AP-2 binding site in the human metallothionein promoter (Imagawa *et al.*, 1987) were used in gel retardation analysis with protein obtained from an *in vitro* translated AP-2 cDNA and nuclear extracts prepared from Jurkat, SLB and HeLa cells. A gel retarded species was seen in Jurkat, SLB and HeLa extracts (Figure 3A, lanes 3–5) that comigrated closely with a species denoted by I obtained from the *in vitro* translated AP-2 cDNA which was not detected in unprogrammed reticulocyte lysate (Figure 3A, lanes 1 and 2). UV crosslinking revealed that the species detected in HeLa, SLB or Jurkat cell lines has a mol. wt of 50 kDa which is close to the predicted mol. wt of AP-2 (data not shown). To demonstrate the specificity of this binding, SLB extract was used in a competition experiment with oligonucleotides corresponding to either a wild-type 21 bp repeat or a 21 bp repeat mutated in the A site. The slow mobility species found in SLB extract that comigrated with *in vitro* translated AP-2 (Figure 3B, lane 1) was completely competed with oligonucleotides corresponding to the wild-type 21 bp III (Figure 3B, lanes 2–4) but not with oligonucleotides lacking these sites (Figure 3B, lanes 5–7). These results suggested that both AP-2 RNA and protein was presented in T-lymphocytes.

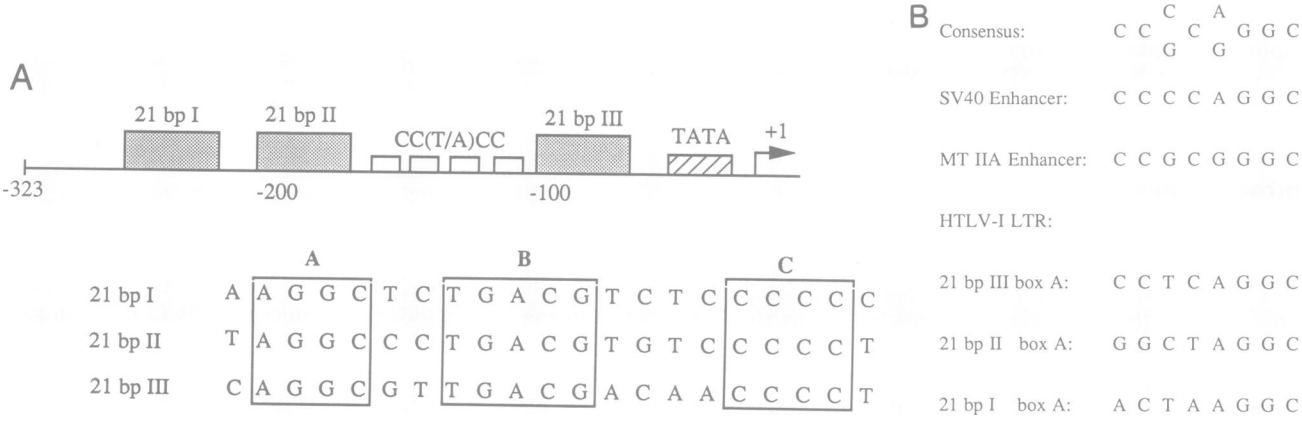

Fig. 1. Consensus binding sequences in the HTLV-I LTR. (**A**) A schematic diagram of the HTLV-I LTR is shown with the positions and sequences of the three 21 bp repeats, the ets binding sites and the TATA box. (**B**) Regions in either the SV40 or metallothionein promoters which bind AP-2 are compared with potential AP-2 binding sites in the HTLV-I 21 bp repeats.

AP-2 binds specifically to the HTLV-I 21 bp repeat A motif

Our studies suggested that AP-2 bound to the A motif in the HTLV-I LTR 21 bp repeat. To elucidate further the role of AP-2 in regulating HTLV-I gene expression, a cDNA encoding the complete AP-2 protein isolated from a HeLa cell λZAP cDNA library (Williams *et al.*, 1988; Williams and Tjian, 1991a,b) was cloned downstream of the glutathione-*S*-transferase (GST) gene in the plasmid pGEX-2T (Smith and Johnson, 1988). This clone made it possible to obtain high level expression of this gene product in *Escherichia coli*. Following affinity purification of this AP-2 fusion protein by glutathione–Sepharose chromatography, both the GST/AP-2 fusion and native AP-2 protein obtained following cleavage of the fusion protein by thrombin were used in gel retardation assays with oligonucleotides corresponding to each of the three HTLV-I 21 bp repeats. The gel retardation results were similar using both authentic AP-2 and the GST/AP-2 fusion protein (data not shown). Only the results with the GST/AP-2 fusions referred to hereafter as AP-2 are shown.

AP-2 (Figure 4, lanes 2, 4 and 6) but not control bacterial extract containing the glutathione-*S*-transferase protein alone (Figure 4, lanes 1, 3 and 5) was capable of binding to each of the 21 bp repeats. The binding of AP-2 was stronger to 21 bp repeat III (Figure 4, lane 6) than for 21 bp repeats I and II (Figure 4, lanes 2 and 4). The 21 bp repeat III has the strongest homology to the SV40 and the metallothionein AP-2 binding sites (Figure 1). To determine the specificity of AP-2 binding, gel retardation with oligonucleotides corresponding to 21 bp repeat III was performed as was competition with both wild-type and mutated unlabeled oligonucleotides (Figure 5). Oligonucleotides corresponding to the wild-type 21 bp repeat competed AP-2 binding (Figure 5, lane 3) as did oligonucleotides containing mutations of either the B or C sites (Figure 5, lanes 5 and 6). However, oligonucleotides containing mutations of the A site either alone (Figure 5,

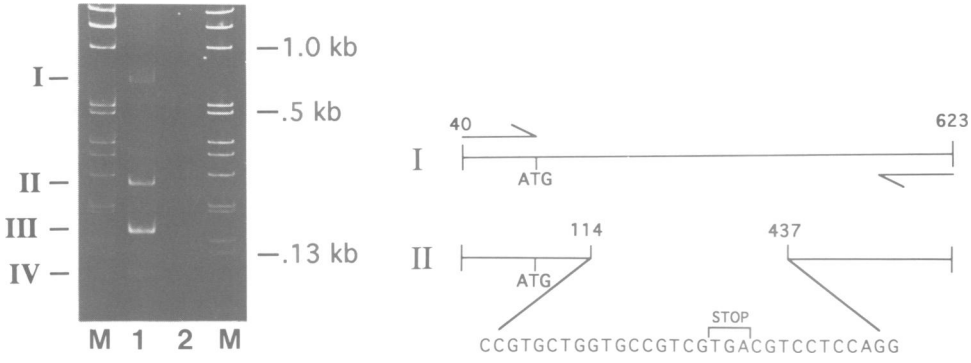

Fig. 2. T-lymphocytes produce AP-2 RNA. PCR are performed on RNA isolated cells from Jurkat cells. The PCR products (I–IV) (lane 1) are shown next to a DNA size standard (M) following gel electrophoresis. A sample of Jurkat RNA which was subject to PCR in the absence of reverse transcription is also shown (lane 2). PCR products were cloned into pUC19 and subjected to DNA sequence analysis. A schematic diagram is shown of the PCR products I and II which each correspond to AP-2 cDNA sequences.

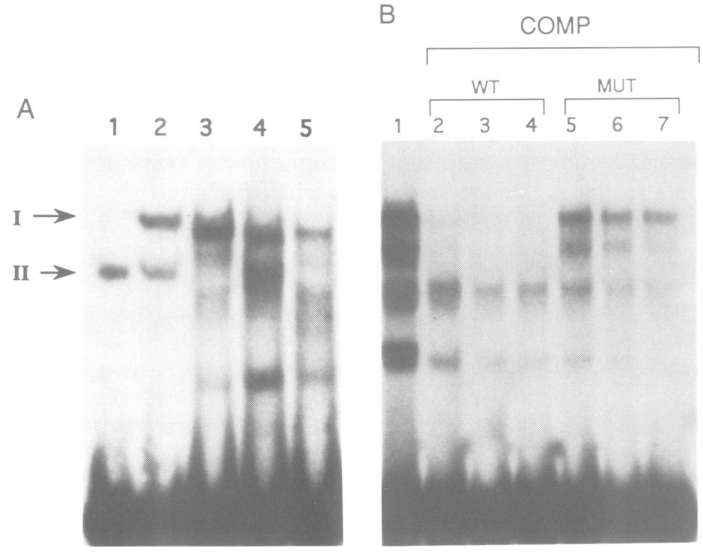

Fig. 3. SLB and Jurkat T-lymphocyte cell lines contain an AP-2 binding activity. Oligonucleotides corresponding to the AP-2 site in the metallothionein promoter were end-labeled and used in gel retardation assays (**A**) with reticulocyte lysate alone (lane 1), reticulocyte lysate programmed with AP-2 mRNA (lane 2) or nuclear extract prepared from SLB (lane 3), Jurkat (lane 4) or HeLa cells (lane 5). Arrow 1 indicates the AP-2 gel retarded species and arrow 2 indicates an AP-2 binding activity found in the reticulocyte lysate. (**B**) Competition using SLB nuclear extract was performed with 0 (lane 1), 50 ng (lane 2), 100 ng (lane 3) or 200 ng (lane 4) of oligonucleotides corresponding to wild-type 21 bp III or 50 ng (lane 5), 100 ng (lane 6) or 200 ng (lane 7) of 21 bp III mutated in the AP-2 binding site.

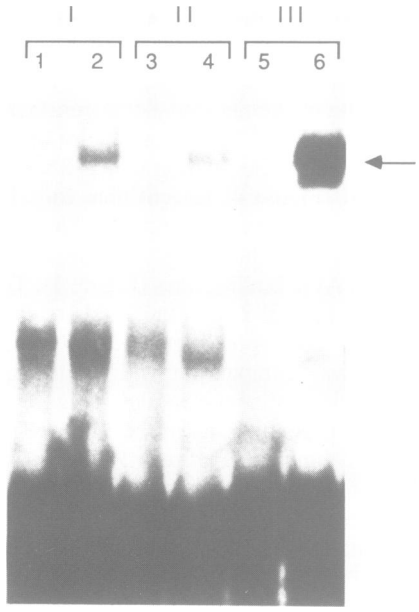

Fig. 4. AP-2 binds to each of the three HTLV-I 21 bp repeats. Oligonucleotides corresponding to HTLV-I 21 bp repeats I (lanes 1 and 2), II (lanes 3 and 4) and III (lanes 5 and 6) were used in gel retardation analysis with bacterial extract containing either glutathione-*S*-transferase (GST) alone (lanes 1, 3 and 5) or GST/AP-2 fusion protein (lanes 2, 4 and 6).

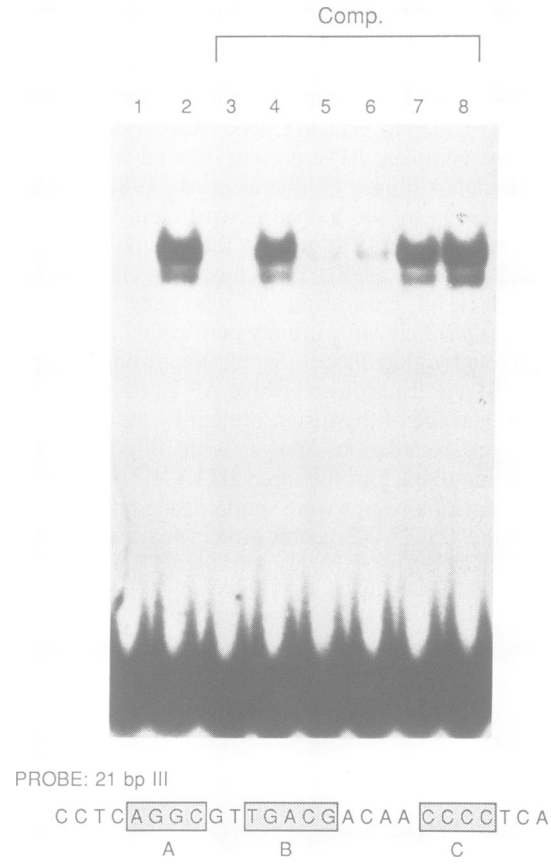

Fig. 5. AP-2 binds to the A motif in the HTLV-I 21 bp repeat. Bacterially produced GST/AP-2 was used in gel regardation assays with oligonucleotides corresponding to the 21 bp III repeat. GST alone (lane 1), GST/AP-2 (lane 2) and GST/AP-2 with a 3-fold excess of unlabeled wild-type oligonucleotides (lane 3) or unlabeled oligonucleotides containing a mutated A site (lane 4), mutated B site (lane 5), mutated C site (lane 6), mutated A and C sites (lane 7) or mutated A, B and C sites (lane 8).

lane 4) or in combination with mutations of either or both of the B and C sites (Figure 5, lanes 7 and 8) did not result in significant competition for AP-2 binding. These results indicate that AP-2 binds specifically to the A site within the 21 bp repeat III. Similar results were seen with 21 bp repeats I and II (data not shown).

Next we determined the specific contact points within the 21 bp repeat required for AP-2 binding. Methylation interference studies were performed using partially methylated oligonucleotides corresponding to either the coding or non-coding strands of the 21 bp repeat III. Following gel retardation analysis with AP-2, the gel retarded species was removed, purified, cleaved with piperidine and electrophoresed, and autoradiography was performed. Methylation of the non-coding strand of the 21 bp repeat revealed that four G residues in the A domain and upstream region were protected (Figure 6A), while on the coding strand two G residues in the A domain were protected (Figure 6B). Free probe cleaved with piperidine gave a uniform cleavage of G residues (data not shown). There was no specific interference to other portions of the 21 bp repeat. These results indicated that AP-2 bound specifically to the A motif in the 21 bp repeats with a strong requirement for the core sequences AGGC found in this motif.

Comparison of AP-2 binding to the HTLV-I 21 bp repeat and metallothionein enhancer

We next wished to determine the binding affinity of AP-2 protein to 21 bp repeat III relative to that seen with the consensus in the AP-2 binding site in the metallothionein promoter. Gel retardation analysis was performed with labeled oligonucleotides corresponding to 21 bp repeat III and either GST (Figure 7, lane 1) or bacterially produced AP-2 (Figure 7, lanes 2 and 8). Competition analysis was performed with unlabeled oligonucleotides corresponding to either the 21 bp repeat III (Figure 7, lanes 3–7) or the metallothionein AP-2 motif (Figure 7, lanes 9–13). The competition analysis with the unlabeled 21 bp repeat oligonucleotides demonstrated that 75% inhibition of AP-2 binding occurred at a 15-fold molar excess of cold competitor. Competition with unlabeled oligonucleotides corresponding to the metallothionein AP-2 binding site demonstrated that 75% inhibition of AP-2 binding was seen with a 3-fold molar excess of cold competitor. These results demonstrated that the binding affinity of AP-2 to 21 bp repeat III was less than but in a range similar to that seen to the metallothionein AP-2 binding site.

AP-2 excludes the binding of a member of the CREB/ATF family to the 21 bp repeat

Within an HTLV-I 21 bp repeat, different cellular factors probably bind to each of the three defined motifs. Therefore it was important to determine whether multiple proteins may bind simultaneously to these different motifs within the 21 bp repeat. We investigated whether AP-2 binding to the A motif could prevent members of the ATF/CREB family such as CREB (Hoeffler *et al.*, 1988) and CRE–BP1 (Maekawa *et al.*, 1989) from binding to the B motif. Gel retardation studies were performed with either uncleaved or a thrombin

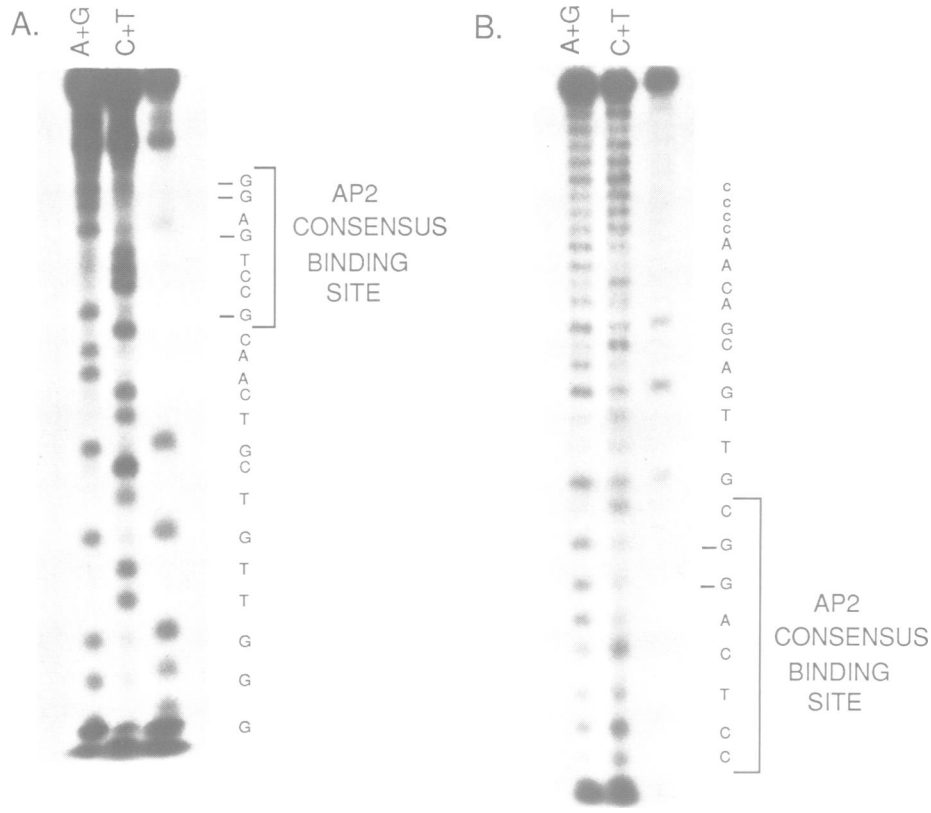

Fig. 6. Methylation interference of AP-2 binding to the 21 bp repeat. Either the non-coding (**A**) or coding (**B**) strands of the HTLV-I 21 bp III were end-labeled and used in methylation interference. Maxam–Gilbert sequencing lanes A+G and C+T, and methylation interference of the GST/AP-2 gel-retarded species are shown. The positions of the G residues in the 21 bp repeat which serve as contacts with AP-2 are indicated by straight lines.

cleaved truncated AP-2 fusion protein designated GST/ΔAP-2 and GST/CRE-BP1. Only the results with the fusion proteins were shown. This truncated form of AP-2 which had binding and dimerization properties that were identical to those of native AP-2 (Williams and Tjian, 1992a,b) was used in gel retardation analysis to separate better the AP2 and CRE-BP1 species. The amount of labeled probe used in the gel retardation resulted in 80–90% saturation so that formation of DNA–protein complexes composed of both AP-2 and CRE-BP1 would be favored.

Gel retardation demonstrated that ΔAP-2 and CRE-BP1 were each capable of binding to oligonucleotides corresponding to 21 bp repeat III (Figure 8, lanes 2 and 3). When equimolar amounts of both proteins were added simultaneously to the labeled oligonucleotides, both ΔAP-2 and CRE-BP1 bound individually to the oligonucleotides (Figure 8, lane 4). No slower mobility species that would indicate the binding of both proteins to the same probe was detected. Thus, both AP-2 and CRE-BP1 could exclude the binding of the other protein to the 21 bp repeat.

Mutant 21 bp repeat III oligonucleotides in either the B motif (Figure 8, lanes 5–8) or the A motif (Figure 8, lanes 9–12) were also used in gel retardation analysis. Mutations in the B motif did not alter ΔAP-2 binding to the 21 bp repeat (Figure 8, lanes 3, 6 and 8) but completely inhibited the binding of CRE-BP1 (Figure 8, lane 7). Likewise, mutations of the A motif eliminated ΔAP-2 binding to the 21 bp repeat (Figure 8, lane 10) but did not affect CRE-BP1 binding (Figure 8, lanes 11 and 12). These results indicated that AP-2 and CRE-BP1 bound independently to different portions of the 21 bp repeat and that the binding of one protein to the 21 bp repeat excluded the binding of the other protein.

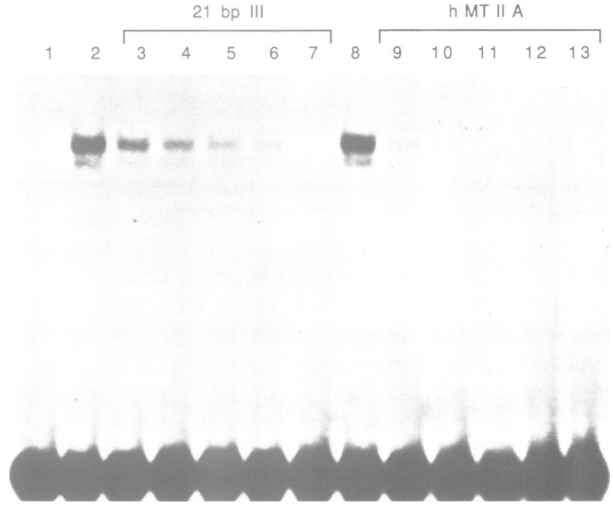

Fig. 7. Binding affinity of AP-2 to the HTLV-I 21 bp III repeat. Labeled oligonucleotides corresponding to the HTLV-I 21 bp repeat III were used in gel retardation analysis with GST alone (lane 1) or GST/AP-2 (lanes 2 and 8). For competition analysis, increasing amounts of unlabeled oligonucleotides corresponding to HTLV-I 21 bp repeat III (lanes 3–7) or the metallothionein AP-2-binding site (lanes 9–13) in either a 10-fold (lanes 3 and 9), 20-fold (lanes 4 and 10), 50-fold (lanes 5 and 11), 100-fold (lanes 6 and 12) or 200-fold (lanes 7 and 13) molar excess were used.

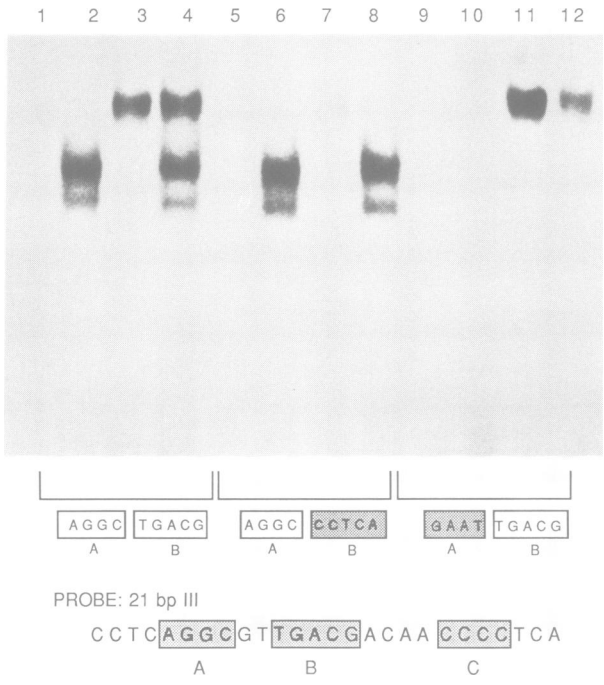

Fig. 8. AP-2 excludes CRE-BP1 binding to the 21 bp repeat. Labeled oligonucleotides corresponding to either 21 bp repeat III (lanes 1–4), 21 bp III with a mutated B site (lanes 5–8) or 21 bp repeat III with a mutated A site (lanes 9–12) were used in gel retardation analysis with GST alone (lanes 1, 5 and 9), in the presence of a truncated GST/ΔAP-2 protein (lanes 2, 6 and 10), in the presence of GST/CRE-BP1 (lanes 3, 7 and 11) or in the presence of both GST/CRE-BP1 and truncated GST/AP-2 proteins (lanes 4, 8 and 12). The position of mutations in either the A or B motifs of the 21 bp repeats is also indicated.

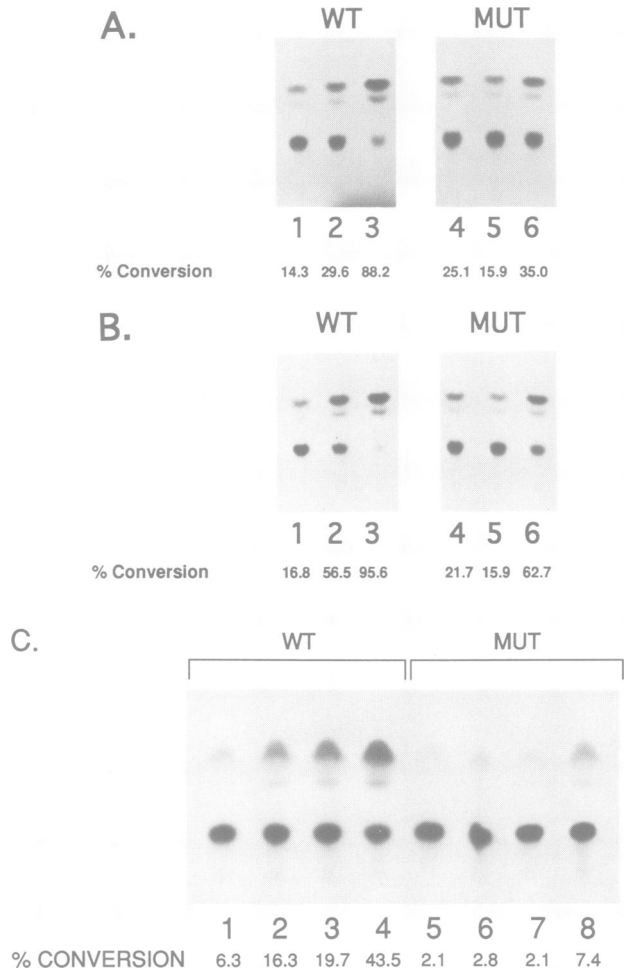

Fig. 9. AP-2 activates the HTLV-I LTR. A wild-type HTLV LTR-I CAT construct (lanes 1–3) or a similar construct mutated in the A site of all three 21 bp repeats (lanes 4–6) were cotransfected into either nonstimulated (**A**) or TPA, PHA and ionomycin stimulated Jurkat cells (**B**) with either expression vector alone (lanes 1 and 4), 1 μg of an AP-2 expression construct (lanes 2 and 5) or 2.5 μg of a tax expression construct (lanes 3 and 6). (**C**) HUT78 cells were transfected with either wild-type (lanes 1–4) or mutated (lanes 5–8) HTLV-I constructs. Lanes 1 and 4 contain no AP-2 expression vector, lanes 2 and 5 contain 1 μg of this expression vector, lanes 3 and 6 contain 3 μg and lanes 4 and 8 contain 2.5 μg of a tax expression vector. The percent CAT conversion for each sample was calculated.

AP-2 activates gene expression from the HTLV-I LTR

Since we demonstrated that AP-2 bound specifically to each of the HTLV-I 21 bp repeats, we next determined the role of this transcription factor in regulating HTLV-I gene expression in T-lymphocytes. Each transfection was repeated three times with similar results. Transfection of an AP-2 expression vector and an HTLV-I LTR CAT construct into unstimulated Jurkat cells resulted in a 2- to 3-fold stimulation of HTLV-I gene expression (Figure 9, lanes 1 and 2). Cotransfection of tax with the HTLV-I LTR CAT construct into unstimulated Jurkat cells resulted in a 6- to 8-fold stimulation of gene expression (Figure 9A, lanes 1 and 3). Mutation of the A sites in each of the three 21 bp repeats in the HTLV-I CAT construct eliminated activation by AP-2 and resulted in only slight stimulation by tax. These results suggested that AP-2 weakly stimulated gene expression of the HTLV-I LTR in unstimulated Jurkat cells. Mutation of the A site decreased both AP-2 and tax activation of the HTLV-I LTR.

Similar transfections were performed with PHA, ionomycin and PMA-stimulated Jurkat cells. AP-2 activated the HTLV-I LTR CAT gene expression 4- to 5-fold in stimulated Jurkat cells (Figure 9B, lane 2) while tax resulted in a 20-fold stimulation of this construct (Figure 9B, lane 3). Mutations of each A site in the HTLV-I LTR eliminated AP-2 activation and decreased tax activation to 3-fold (Figure 9B, Lanes 4–6). Four- to 5-fold activation of HTLV-I gene expression by AP-2 was also seen when these constructs were transfected into HUT78 (Figure 9C) and SLB cells (data not shown). These results suggested that both AP-2 and tax were capable of significant activation of HTLV-I gene expression in T-lymphocytes.

Next we defined the number of 21 bp repeats required for activation by AP-2 and tax. A schematic of the constructs used in these studies is shown in Figure 10. All transfections were performed in four independent experiments on both Jurkat and HepG2 cells, a cell line which has been demonstrated to contain non-detectable levels of AP-2 (Williams *et al.*, 1988). The result obtained for each construct is given in fold activation. Significant activation of the HTLV-I LTR by AP-2 expression vectors required at least two 21 bp repeats (Figure 10, lines 2 and 3). HTLV-I CAT constructs containing either 21 bp repeats II and III or I and II gave levels of gene expression with AP-2 that were slightly decreased as compared with that seen with all three 21 bp repeats (Figure 10, lines 1–3). All these

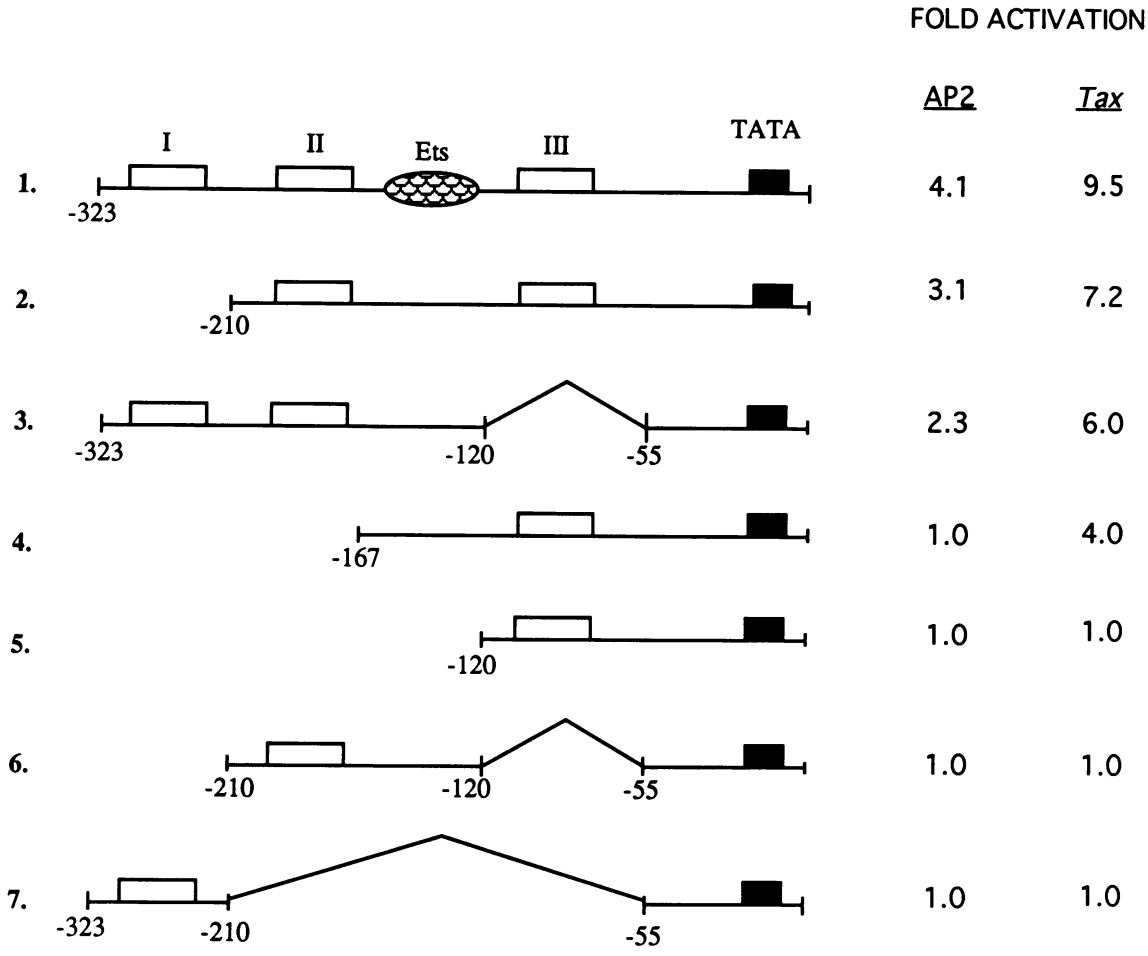

Fig. 10. Activation of HTLV-I mutants by AP-2 and tax. A variety of truncated HTLV-I LTR constructs were cotransfected into Hep G2 cells with either an AP-2 expression vector (1 μg), a tax expression vector (2.5 μg) or in the absence of these factors. The final concentration of expression vector in each transfection was kept constant at 10 μg by adding RSV β-globin. The results are the summary of four independent experiments and the results are expressed in fold of activation of each construct by AP-2 and tax.

constructs were also strongly activated by tax, as previously reported by Fujisawa *et al.* (1986). Constructs containing only one of the 21 bp repeats were not significantly activated by AP-2 (Figure 10, lines 4–7). These constructs were also poorly activated by tax except for one construct which contained 21 bp repeat III and an element which binds both the ets (Bosselut *et al.*, 1990; Gitlin *et al.*, 1991) and SP1 (Nyborg *et al.*, 1990) proteins (Figure 10, line 4).

Discussion

Multiple regions in the HTLV-I LTR are critical for modulating gene expression (Fujisawa *et al.*, 1986, 1989; Paskalis *et al.*, 1986; Shimotohno *et al.*, 1986; Brady *et al.*, 1987; Rosen *et al.*, 1987; Bosselut *et al.*, 1990; Montagne *et al.*, 1990). These regulatory regions serve as the binding sites for a variety of cellular proteins (Altman *et al.*, 1988; Jeang *et al.*, 1988; Nyborg *et al.*, 1988, 1990; Tan *et al.*, 1989; Bosselut *et al.*, 1990; Montagne *et al.*, 1990; Nyborg and Dynan, 1990; Yoshimura *et al.*, 1990; Béraud *et al.*, 1991). The transactivator protein, tax, is also required for high levels of HTLV-I gene expression (Sodroski *et al.*, 1984; Chen *et al.*, 1985; Felber *et al.*, 1985; Fujisawa *et al.*, 1985; Smith and Greene, 1990). In this study, we investigate the role of the transcription factor AP-2 on the regulation of HTLV-I gene expression.

Fractionation of cellular extracts reveals that multiple factors can bind to the 21 bp repeat (Altman *et al.*, 1988; Jeang *et al.*, 1988; Nyborg *et al.*, 1988, 1990; Tan *et al.*, 1989; Bosselut *et al.*, 1990; Montagne *et al.*, 1990; Nyborg and Dynan, 1990; Yoshimura *et al.*, 1990). These factors appear in some instances to have cell-type specificity though most factors have not been completely characterized (Nyborg and Dynan, 1990). One study reports that oligonucleotide affinity column purified AP-2 from HeLa cells binds to a region of the HTLV-I LTR which overlaps each of the 21 bp repeats, but AP-2 was not detected in SLB cells using DNase I footprinting assays (Nyborg and Dynan, 1990). However, we detected AP-2 RNA and binding activity in several T-lymphocyte cell lines using PCR and gel retardation assays. The level of AP-2 RNA and protein in HeLa cells is also relatively low since AP-2 expression vectors activate constructs with AP-2 binding sites in transfection assays performed in HeLa cells (Williams *et al.*, 1988). Other studies, using λgt11 expression cloning with oligonucleotides similar to those found in the B motif, demonstrate that multiple members of the CREB/ATF family bind to this region (Hai *et al.*, 1989; Tsujimoto *et al.*, 1991; Yoshimura *et al.*, 1990). Which if any of these factors is/are responsible for tax activation has not yet been demonstrated. In addition, constructs containing isolated CREB/ATF sites are not inducible by tax, suggesting that either the position of these

sites in a specific promoter or proteins binding to overlapping motifs are critical for tax activation (Montagne *et al.*, 1990). The fact that both the A and C sites in the 21 bp repeat are critical for tax activation suggests that multiple factors may be required for tax activation.

Since the A motif in the 21 bp repeat has strong homology with AP-2 binding sites found in the SV40 and metallothionein promoters, we wished to determine whether AP-2 could bind and activate gene expression from the HTLV-I LTR 21 bp repeats. AP-2 RNA is expressed in T-lymphocytes and a cellular protein of similar molecular weight and mobility to AP-2 binds to the A motif in the 21 bp repeat. To investigate AP-2 binding to the HTLV-I LTR further, we demonstrate that protein expressed from an AP-2 cDNA binds specifically and with high affinity to the A motif in the 21 bp repeat. The binding of AP-2 to the 21 bp repeat excludes the binding of a member of the ATF/CREB family, CRE-BP1 (Maekawa *et al.*, 1989), to an adjacent binding site in the B motif. Finally, we demonstrate that AP-2 can activate expression of the HTLV-I LTR gene in both T-lymphocytes and other cell lines. Thus AP-2 may be important in regulating HTLV-I gene expression during infection of both lymphoid and non-lymphoid cells (Clapham *et al.*, 1983).

AP-2 binds and activates gene expression from other viral promoters, including SV40 (Mitchell *et al.*, 1987; Williams *et al.*, 1988) and hepatitis B (Seto *et al.*, 1990). Within different viral and cellular promoters there appears to be marked degeneracy of the AP-2 consensus binding sequences. The binding domain of AP-2 has no detectable homology with other well described DNA binding proteins. In addition, AP-2 contains a carboxyl-terminal dimerization domain and an amino-terminal proline-rich activating domain (Williams and Tjian, 1991a,b). UV crosslinking experiments indicate that AP-2 binds to the HTLV-I LTR as a dimer, in agreement with its binding to other promoters (data not shown, Williams and Tjian, 1991a,b). The finding that AP-2 excludes binding of members of the CREB/ATF family to the B motif suggests that different factors may compete for the binding to the 21 bp repeat resulting in different patterns of gene expression. Since the activity of AP-2 is regulated by both phorbol esters and cyclic AMP (Imagawa *et al.*, 1987; Luscher *et al.*, 1989), AP-2 may mediate expression of HTLV-I by various stimuli which affect the signal transduction pathway (Poteat *et al.*, 1989; Tan *et al.*, 1989b). Thus, HTLV-I gene expression may be regulated by cellular factors such as the recently described HEB1 (Béraud *et al.*, 1991) which potentially activate gene expression in a tax-dependent manner, in addition to other factors which function independent of tax.

The role of tax in modulating AP-2 activation of HTLV-I gene expression is not addressed in this study. However, experiments suggest that tax inhibits AP-2 activation of HTLV-I gene expression by direct interaction of tax with AP-2. This situation is similar to the interaction of SV40 T antigen with AP-2 which results in inhibition of SV40 gene expression. In addition, tax has also been demonstrated to regulate negatively DNA polymerase β which contains potential AP-2 binding sites (Mitchell *et al.*, 1987; Jeang *et al.*, 1990). Thus, one possible mechanism of tax regulation may be to inhibit the binding of one group of cellular transcription factors and allow the binding of different factors which may result in altered patterns of gene expression. Further studies will be required to define the role of AP-2 in activating HTLV-I gene expression in both the presence and absence of tax and its role in the pathogenesis of HTLV-I infection.

Materials and methods

Plasmid constructs

A HeLa cDNA library cloned into λZAP was screened with 30mer oligonucleotides corresponding to both the amino- and carboxyl-termini of the AP-2 coding sequences. An AP-2 cDNA spanning the entire AP-2 coding region extending from nt 8 to 1621 was isolated (Williams *et al.*, 1988). This cDNA was cloned in pGEM2, downstream of an SP6 promoter for *in vitro* expression. A *Bam*HI site was inserted into the amino-terminus of AP-2 and a *Bam*HI−*Eco*RI fragment from this construct was cloned into pGEX-2T (Pharmacia) for bacterial expression (Smith and Johnson, 1988). A truncation of AP-2 spanning amino acids 165−437 was also cloned into the *Sma*I and *Eco*RI sites of pGEX-2T for bacterial expression. A cDNA containing the complete coding sequence of CRE-BP1 (Maekawa *et al.*, 1988) was isolated from a λgt11 library. A *Bam*HI site was inserted upstream of the ATG using PCR and a *Bam*−*Eco*RI fragment was cloned into pGEX-2T. For transfection experiments, an *Eco*RI fragment spanning the AP-2 cDNA was cloned into the *Eco*RI site of pDP18 (Garcia *et al.*, 1989). A tax cDNA (Wano *et al.*, 1988) was cloned into the *Eco*RI site of pDP18.

PCR reactions

AP-2 cDNA was hexamer primed from 10 μg of total Jurkat E cDNA using reagents and conditions prescribed in a Promega Riboclone cDNA synthesis kit. A mock cDNA synthesis was performed by deleting AMV reverse transcriptase from the reaction mix and was used as a negative control template for subsequent PCR analysis. One-tenth of the hexamer primed cDNA was used as a template for PCR which was performed in a DNA thermal cycler using reagents from a GeneAmp Kit. The reaction was heated for 5 min at 94°C and was subsequently incubated for 40 cycles of PCR (55°C, 1 min; 72°C, 1 min; 94°C, 1 min). 2 μl of the reaction was reamplified for 40 cycles using a nested antisense primer. 25mer oligonucleotide primers used for PCR contained sequences corresponding to the following positions in the AP-2 cDNA (Williams *et al.*, 1988): sense primer (nt 63−87), 5′-ATGCTTTTGGAAATTGGACGGGATAATA-3′; antisense primer (nt 701−725), 5′-GAGGGAGCGAGAGGCGACCCG-GAACT-3′; antisense nested primer (nt 599−623), 5′-GCTGTTGGAC-TTGGACAGGGACACG-3′. PCR products were separated on a 5% polyacrylamide gel, electroeluted from the polyacrylamide, phosphorylated with T4 polynucleotide kinase and cloned into pUC19 for DNA sequencing.

In vitro transcription and translation

For *in vitro* translation, AP-2/pGEM2 was linearized with *Hin*dIII and transcribed *in vitro* using a commercial kit for RNA synthesis (Promega). RNA was translated *in vitro* with micrococcal nuclease treated rabbit reticulocyte lysate in 50 μl reaction volume, as directed by the supplier (Promega). For analysis of the translation products, proteins were labeled with [^{35}S]methionine and a 5 μl aliquot was run on an 8% SDS−polyacrylamide gel. After electrophoresis, the gel was dried and subject to autoradiography.

Bacterial expression of AP-2 and CRE-BP1

CRE-BPI/pGEX-2T, AP-2/pGEM-2T and ΔAP-2/pGEX-2T constructs and a control pGEX-2T without insert were transformed into *E.coli* strain JM109. 500 ml cultures of *E.coli* were grown to an OD_{600} of 0.7−0.9 and induced with 0.4 mM IPTG for 3 h. Cells were pelleted and resuspended in 10 ml PBS with 1% Triton X-100 (Sigma). Cells were mildly sonicated, debris was pelleted and supernatant incubated with 1 ml glutathione−Sepharose for 90 min at 4°C. The beads were then washed five times with 10 ml of PBS. Elution was done in 3 ml Tris−HCl, pH 8.0, 10 mM glutathione incubated for 90 min at 4°C. Protein was dialyzed against protein storage buffer (10 mM Tris−HCl, pH 8.0, 20% glycerol, then against 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF).

Gel retardation analysis

For gel retardation analysis 1−2 μg of nuclear extract (Dignam *et al.*, 1983) prepared from SLB, Jurkat or HeLa cell lines was used. Gel retardation with *in vitro* translated AP-2 was done with 1 μl of a 50 μl reaction. When bacterially produced AP-2 (Williams *et al.*, 1988) or CRE-BP1 (Maekawa *et al.*, 1988) proteins (either fused to glutathione-S-transferase or cleaved with thrombin) were used, 50−100 ng of extract was added to each reaction. In all cases, extracts were incubated with 1 μg of poly(dI-dC) (Pharmacia) in binding buffer (10 mM Tris−HCl, pH 7.4, 1 mM EDTA, 1 mM DTT, 5% glycerol, BSA 50 μg/ml, 50 mM NaCl) for 10 min. When competition

was performed, the proteins were incubated for a further 10 min with 50 ng cold oligonucleotide unless otherwise indicated. The labeled probe (50 000 c.p.m.) was added to the mixture and after 20 min, incubation at room temperature, the samples were loaded on a 4% polyacrylamide gel and electrophoresed in 0.25 × TBE (1 × TBE is 89 mM Tris/89 mM borate pH 8.3, 1 mM EDTA). Oligonucleotides used had the following sequences: 21 bp repeat I, 5′-CCAGACTAAGGCTCTGACGTCTCCCCCC-CGGA-3′; 21 bp repeat II, 5′-CTCGGGCTAGGCCCTGACGTG-TCCCCCTGAA-3′; 21 bp repeat III, 5′-TCGACGTCCTCAGGC-GTTGACGACAACCCTCAG-3′; mutated A box 21 bp III, 5′-TCG-ACGTCCTCGAATGTTGACGACAACCCCTCAC-3′; mutated B box 21 bp III, 5′-TCGACGTCCTCAGGCGTCAATAACAACCCCTCAC-3′; mutated C box 21 bp III, 5′-TCGACGTCCTCAGGCGTTGACGA-CAATTATTCAC-3′; mutated A and C boxes III, 5′-TCGACGTC-CTCGAATGTTGACGACAATTATTCAC-3′; mutated A, B, and C boxes III 5′-CGACGTCCTCGAATGTCAATAACAATTATTCAC-3′; metallothionein 5′-TCGACTGACCGCCCGCGGCCCGTTGTGCAGC-3′.

HTLV-I LTR CAT deletion constructs

All constructs containing the HTLV-I LTR were inserted into the pUC19 polylinker of pJGFCAT (Garcia *et al.*, 1989). *In vitro* site-directed mutagenesis (Amersham) was performed to introduce the identical mutations into the A box as shown in the A box mutated oligonucleotides used in gel retardation analysis. An additional site-directed mutation (T→C at position −120) was introduced for cloning purposes.

The wild-type HTLV-I LTR construct was constructed by inserting a *Sma*I (−323)−*Hha*I (+24) end-filled fragment into the *Sma*I site in pJGFCAT. The −210 construct contained an end-filled *Ava*I−*Hha*I fragment; the −167 construct contained an *Alu*I−*Hla*I end-filled fragment; and the −120 fragment contained an *Nla*III−*Hha*I end-filled fragment. These fragments were all inserted into pJGFCAT 19 restricted with *Sma*I or *Sph*I−*Sma*I. The −210 Δ−120/−55 construct contained the end-filled *Ava*I−*Nla*III fragment joined to an *Nde*I(−55)/*Hha*I(+24) end-filled fragment and the −323 Δ−210/−55 fragment contained a *Sma*I(−323)−*Ava*I(−210) fragment joined to an *Nde*I(−55)/*Hha*I(+24) end-filled fragment. Each construct was cloned into the *Sma*I site of pJGFCAT 19.

Transfections and CAT assays

Jurkat, HUT78 and SLB cells were maintained in RPMI and 10% FBS with penicillin and streptomycin. Cells were diluted 1:3 in fresh medium on the day prior to transfection. 1.5 × 10^7 cells were transfected by electroporation in a volume of 250 μl using a capacity of 1180 μF loaded at 250 V. Cells were then incubated for 16 h. Half of the sample was then transferred to a new flask and induced with TPA (50 ng/ml), PHA (4 μg/ml) and ionomycin (2 μM). Cells were harvested after 8 h and one-quarter of the extract was used for CAT assays (Gorman *et al.*, 1982). HepG2 cells were maintained in DMEM with penecillin and streptomycin and 10% FBS. Cultures were split on the day prior to transfection and used at 40−50% confluency. Transfection by calcium phosphate precipitation, and CAT assays performed as described (Gorman *et al.*, 1982).

Acknowledgements

We wish to thank Warner Greene for supplying a tax cDNA, Dixie Shurland for technical assistance and Chuck Leavitt for the preparation of this manuscript. This work was supported by grants from the National Institute of Health, Veterans Administration, and UC AIDS Task Force. R.G. was supported by a grant from an American Cancer Society Faculty Research Award and C.M. from a NATO grant.

References

Altman,R., Harrich,D., Garcia,J.A. and Gaynor,R.B. (1988) *J. Virol.*, **62**, 1339−1346.

Béraud,C., Lombard-Platet,G., Michal,Y. and Jalinot,P. (1991) *EMBO J.*, **10**, 3795−3803.

Bosselut,R., Duvall,J.F., Gégonne,A., Bailly,M., Hémar,A., Brady,J. and Ghysdael,J. (1990) *EMBO J.*, **9**, 3137−3144.

Brady,J., Jeang,K.T., Duvall,J. and Khoury,G. (1987) *J. Virol.*, **61**, 2175−2181.

Chen,I.S.Y., Slamon,D.J., Rosenblatt,D.J., Shah,N.P., Queen,S.G. and Wachsman,W. (1985) *Science*, **229**, 54−58.

Clapham,P., Nagy,K., Cheimgsong-Popov,R. and Weiss,R. (1983) *Science*, **222**, 1125−1127.

Dignam,J.D., Lebowitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475−1489.

Felber,B.K., Paskalis,H., Kleinmann-Ewing,C., Wong-Staal,F. and Pavlakis,G.N. (1985) *Science*, **229**, 675−679.

Fujisawa,J., Seiki,M., Kiyokawa,T. and Yoshida,M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2277−2281.

Fujisawa,J., Seiki,M., Seto,M. and Yoshida,M. (1986) *EMBO J.*, **5**, 713−718.

Fujisawa,J., Toita,M. and Yoshida,M. (1989) *J. Virol.*, **63**, 3234−3239.

Garcia,J.A., Harrich,D., Soultanakis,E., Wu,F., Mitsuyasu,R. and Gaynor,R.B. (1989) *EMBO J.*, **8**, 765−778.

Gessain,A., Barin,F., Vernant,J.C., Gout,O., Maurs,L., Calender,A. and de The,G. (1985) *Lancet*, **ii**, 407−410.

Gitlin,S., Bosselut,R., Gegonne,A., Ghysdael,J. and Brady,J.N. (1991) *J. Virol.*, **65**, 5513−5523.

Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044−1051.

Hai,T., Liu,F., Coulo,W. and Green,M. (1989) *Genes Dev.*, **3**, 2083−2090.

Hinrichs,S.H., Nerenberg,M., Reynolds,S.R.K., Khoury,G. and Jay,G. (1987) *Science*, **237**, 1340−1343.

Hoeffler,J.P., Meyer,T.E., Yun,Y., Jameson,J.L. and Haebner,J.F. (1988) *Science*, **242**, 1430−1433.

Imagawa,M., Chiu,R. and Karin,M. (1987) *Cell*, **51**, 251−260.

Jeang,K.T., Boros,I., Brady,J., Radonovich,M. and Khoury,G. (1988) *J. Virol.*, **62**, 4499−4509.

Jeang,K.T., Widen,S.G., Semmes,O.J. and Wilson,S.H. (1990) *Science*, **247**, 1082−1084.

Koeffler,H.P., Chen,I.S.Y. and Golde,D.W. (1984) *Blood*, **64**, 482−490.

Luscher,B., Mitchell,P.J., Williams,T. and Tjian,R. (1989) *Genes Dev.*, **3**, 1507−1517.

Maekawa,T., Sakura,H., Kaneii-Ishii,C., Sudo,T., Yoshimura,T., Fujisawa,J., Yoshida,M. and Ishii,S. (1989) *EMBO J.*, **8**, 2023−2028.

Marriot,S.J., Boros,I., Duvall,J.F. and Brady,J.N. (1989) *Mol. Cell. Biol.*, **9**, 4152−4160.

Marriot,S.J., Lindholm,P.F., Brown,K.M., Gitlin,S.D., Duvall,J.F., Radonovich,M.F. and Brady,J.N. (1990) *Mol. Cell. Biol.*, **10**, 4192−4202.

Mitchell,P.J., Wang,C. and Tjian,R. (1987) *Cell*, **50**, 847−861.

Montagne,J., Béraud,C., Crenon,I., Lombard-Platet,G., Gazolo,L., Sargent,A. and Jalinot,P. (1990) *EMBO J.*, **9**, 957−964.

Nyborg,J.K. and Dynan,W.S. (1990) *J. Biol. Chem.*, **265**, 8230−8236.

Nyborg,J.K., Dynan,W.S., Chen,I.S.Y. and Wachsman,W. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1457−1461.

Nyborg,J.K., Matthews,M.A., Yucel,J., Walls,L., Golde,W.T., Dynan,W.S. and Wachsman,W. (1990) *J. Biol. Chem.*, **265**, 8227−8242.

Osame,M., Usuku,K., Ijichi,N., Amitani,H., Igata,A., Matsumoto,M. and Tara,H. (1986) *Lancet*, **i**, 1031−1032.

Paskalis,H., Felber,B.K. and Pavlakis,G.N. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 6558−6562.

Poiesz,B.F., Ruscetti,F.W., Gazdar,A.F., Bunn,P.A., Minna,J.D. and Gallo,R.C. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7415−7419.

Poiesz,B.F., Ruscetti,F.W., Reitz,M.S., Kalyanaraman,V.S. and Gallo,R.C. (1981) *Nature*, **294**, 268−271.

Poteat,H.T., Kadison,P., McGuire,K., Park,L., Park,R.E., Sodrowski,J.G. and Haseltine,W.A. (1989) *J. Virol.*, **63**, 1604−1611.

Pozzatti,R., Vogel,J. and Jay,G. (1990) *Mol. Cell. Biol.*, **10**, 413−417.

Rosen,C.A., Park,R., Sodroski,J.G. and Heseltine,W.A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4919−4923.

Seto,E., Mitchell,P.J. and Yen,T.S.B. (1990) *Nature*, **344**, 72−74.

Shimotohno,K., Takano,M., Terunchi,T. and Miwa,M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8112−8116.

Smith,M.R. and Greene,W.C. (1990) *Genes Dev.*, **4**, 1875−1885.

Smith,D.B. and Johnson,K.S. (1988) *Genes Dev.*, **67**, 31−40.

Sodroski,J.G., Rosen,C., Goh,W.C. and Haseltine,W. (1984) *Science*, **225**, 381−385.

Tan,T.H., Horikoshi,M. and Roeder,R.G. (1989) *Mol. Cell. Biol.*, **9**, 1733−1745.

Tanaka,A., Takahashi,C., Yamaoka,S., Nosaka,T., Maki,M. and Hatanaka,M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1071−1075.

Tsujimoto,A., Nyunoya,H., Morita,T., Seto,T. and Shimotohno,K. (1991) *J. Virol.*, **65**, 1420−1426.

Wano,Y., Feinberg,M., Hosking,J.B., Bogerd,H. and Greene,W.C. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 9733−9737.

Williams,T. and Tjian,R. (1991a) *Genes Dev.*, **5**, 670−682.

Williams,T. and Tjian,R. (1991b) *Science*, **251**, 1067−1071.

Williams,T., Admon,A., Luscher,B. and Tjian,R. (1988) *Genes Dev.*, **2**, 1557−1569.

Yoshida,M., Miyoshi,I. and Hinuma,Y. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2031−2035.

Yoshimura,T., Fujisawa,J.I. and Yoshida,M. (1990) *EMBO J.*, **9**, 2537−2542.

Received on December 13, 1991; revised on March 24, 1992