JOURNAL OF VIROLOGY, July 1994, p. 4597–4608
0022-538X/94/$04.00+0

Vol. 68, No. 7

Identification of Human T-Cell Lymphotropic Virus Type I 21-Base-Pair Repeat-Specific and Glial Cell-Specific DNA-Protein Complexes

MARIBETH TILLMANN, RENEE WESSNER, AND BRIAN WIGDAHL*

Department of Microbiology and Immunology, College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania 17033

Received 7 October 1993/Accepted 20 April 1994

The human T-cell lymphotropic virus type I (HTLV-I)-encoded protein, Tax, is capable of *trans*-activating HTLV-I transcription by interacting with specific sequences in the HTLV-I long terminal repeat (LTR) which comprise an inducible enhancer containing three imperfect tandem repeats of a 21-bp sequence. There is no evidence that purified Tax can bind to DNA in the absence of cellular factors, suggesting that Tax most likely regulates transcription via interaction with cellular factors. Since HTLV-I is a documented agent of adult T-cell leukemia and tropical spastic paraparesis, disorders of the immune and nervous systems, respectively, characterization of cellular factors of lymphoid and neuroglial origin which interact with the 21-bp repeat elements is essential to understanding of the mechanisms involved in basal and Tax-mediated transcription in cells of immune and nervous system origin. Utilizing electrophoretic mobility shift (EMS) analyses, we have detected both 21-bp repeat-specific and glial cell-specific DNA-protein complexes. Several 21-bp repeat-specific DNA-protein complexes were detected when nuclear extracts derived from cells of lymphoid (Jurkat, SupT1, and H9), neuronal (IMR-32 and SK-N-MC), and glial (U-373 MG, Hs683, and U-118) origin were used in reactions with each of the three 21-bp repeat elements. In addition, a glial cell-specific DNA-protein complex was detected when nuclear extracts derived from U-373 MG, Hs683, and U-118 glial cell lines reacted with the promoter-distal and central 21-bp repeat elements. Furthermore, EMS analyses performed with nuclear extracts derived from lymphocytic and glial cell origin and a 223-bp fragment of the HTLV-I long terminal repeat encompassing the three 21-bp repeat elements (designated Tax-responsive elements 1 and 2, TRE-1/-2) have also resulted in the detection of glial cell type-specific DNA-protein complexes. Competition EMS analyses with oligonucleotides containing transcription factor binding site sequences indicate the involvement of a cyclic AMP response element binding protein in the formation of DNA-protein complexes which form with all three 21-bp repeat elements and the glial cell-specific DNA-protein complex as well as the involvement of Sp1 or an Sp1-related factor in the formation of the 21-bp repeat III-specific DNA-protein complexes.

Human T-cell lymphotropic virus type I (HTLV-I) is a documented etiologic agent of both hematologic and neurologic disorders. Specifically, HTLV-I has been determined to be an etiologic agent of both adult T-cell leukemia and a slowly progressive neurologic disorder, tropical spastic paraparesis (3, 18, 40, 46, 50, 51). However, the viral and host factors involved in determining the ultimate outcome of HTLV-I infection (malignancy versus neurologic dysfunction) remain to be elucidated. While the interaction of HTLV-I with cells of lymphoid origin has been examined extensively, relatively little is known concerning the interaction of HTLV-I with cells residing in the central nervous system.

Any aspect of the retroviral life cycle, including entry, reverse transcription, integration, transcription, and assembly, may impact on the oncogenic and/or neuropathogenic processes associated with virus infection. Furthermore, these processes may be highly dependent on interactions occurring between retrovirus- and host cell-specific components. Subsequent to viral entry, the outcome of HTLV-I infection in a given cell type within either the immune or nervous system is critically dependent on cellular factors that interact with retroviral sequences involved in transcriptional regulation. The integrated retroviral genome is flanked at both ends by noncoding long terminal repeat (LTR) sequences composed of three regions, U3, R, and U5 (Fig. 1). The LTRs are integral components of the viral regulatory system containing information essential to the regulation of retroviral integration, transcription, and replication.

Evidence demonstrating that the LTR sequences of some retroviruses play a role in tissue and cell type specificity and may also be involved in determining the course of disease associated with retroviral infection has been presented (7, 8, 31, 41). Use of the murine retrovirus system has suggested that the in vitro host range of selected murine leukemia viruses is determined by the retroviral LTR, specifically the U3 region (7, 8, 31, 41). For several murine retroviruses, minor variations in the proviral genome, particularly the LTR or envelope gene, result in alterations in the cellular tropism of the retrovirus and, in turn, in its pathogenicity (7, 31, 37, 41). In addition, the promoters of several retroviruses, including Moloney and Friend murine leukemia viruses and human immunodeficiency virus, have been implicated in cell type-specific expression (7–9, 31, 41). For example, when transgenic mice were generated by utilizing the LTRs from either central nervous system-derived or T-cell-derived human immunodeficiency virus strains, expression of the reporter gene was detected only in mice transgenic for the central nervous system-derived LTRs (9). Furthermore, in support of the idea that the LTR has a role in tissue type- or cell type-specific expression of HTLV-I, the generation of transgenic mice bearing a transgene composed of the HTLV-I LTR isolated from a patient with tropical

* Corresponding author. Fax: (717) 531-5580. Electronic mail address: bwigdahl@cor-mail.biochem.hmc.psu.edu.

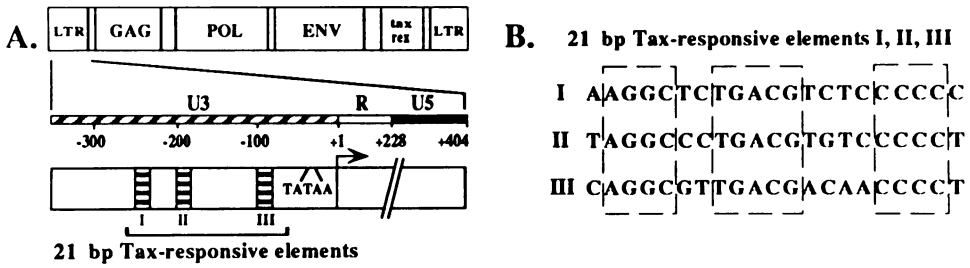

FIG. 1. (A) Structure of the HTLV-I LTR in the context of the viral genome. The 21-bp repeat elements are located within the U3 region of the LTR at positions −251 to −231, −203 to −183, and −103 to −83 relative to the transcriptional start site. (B) Sequence comparison of the three 21-bp repeat elements. The regions of strict conservation among the 21-bp repeat elements are indicated by the broken boxes representing (from left to right) domains A, B, and C.

spastic paraparesis linked to a reporter gene demonstrated that LTR-directed expression of the reporter gene occurred primarily in cells of the central nervous system (20).

With respect to HTLV-I, virus-specific transcription is dependent on *cis*-acting enhancer sequences comprising three 21-bp imperfectly repeated elements located within the U3 region of the HTLV-I LTR at positions −251 to −231, −203 to −183, and −103 to −83 relative to the transcriptional start site (Fig. 1) (38, 43). Immediately following virus infection, low but necessary levels of HTLV-I transcription which are critically dependent on cellular transcription factors are attained. Basal levels of transcription are greatly enhanced by the virus-encoded regulatory protein, Tax, which mediates *trans* activation of the HTLV-I LTR through the 21-bp repeat elements (14, 42, 44, 46, 47). As with basal transcription, Tax-mediated transcriptional *trans* activation of the HTLV-I LTR requires the participation of cellular intermediaries (16, 19, 38, 46, 48, 53). Mutagenesis studies have shown that at least two 21-bp repeat elements are necessary for efficient transcriptional activation (14, 39, 45, 46). A number of proteins, of both cellular and viral origin, which interact either directly or indirectly with sequences in the HTLV-I LTR have been identified (see references 33 and 46 for reviews). However, the proteins in these studies were identified by utilizing cell lines derived from sources other than the nervous system. Because the interaction of host cellular factors with the 21-bp repeat elements and possibly other regions of the HTLV-I LTR is critical for basal and Tax-mediated expression, further identification and characterization of the cellular factors of immune system and nervous system origins which interact with the LTR regulatory unit is essential to understanding of the molecular mechanisms involved in HTLV-I LTR-directed transcription in these cell populations.

While each of the 21-bp repeat elements has three strictly conserved domains, termed A, B, and C (Fig. 1) (46), these domains comprise only 13 of the 21 base pairs. Studies have suggested, utilizing the DNase I protection assay, that there is differential binding of cellular factors to each of the three 21-bp elements (36). Differential binding of cellular factors to the enhancer elements, in turn, may play a role in basal and Tax-mediated HTLV-I LTR-directed transcription. Therefore, characterization of the DNA-protein complexes formed between cellular factors and each of the 21-bp repeat elements is warranted. In addition, identification of tissue type- or cell type-specific factors which interact with these elements may provide important information concerning the regulation of HTLV-I gene expression in cells of immune versus nervous system origin. In an effort to characterize the interaction of cellular factors with each of the three 21-bp repeat elements, we have utilized double-stranded oligonucleotides homologous to each of the 21-bp repeat elements as well as a native fragment of the HTLV-I LTR containing the three 21-bp repeat elements (designated Tax-responsive elements 1 and 2, TRE-1/-2) in reactions with nuclear extracts derived from selected cell lines of lymphocytic, neuronal, and glial origins in electrophoretic mobility shift (EMS) analyses. Both 21-bp repeat-specific DNA-protein complexes and a glial cell type-specific DNA-protein complex were detected, whose formation may involve Sp1 or an Sp1-related factor and a cyclic AMP response element (CRE) binding protein (CREB)-activating transcription factor (ATF) family member(s), respectively. These studies also indicate that the three 21-bp repeat elements are not functionally equivalent with respect to DNA-protein complex formation.

MATERIALS AND METHODS

Cells. Human T-cell lymphocytic cell lines utilized in these studies include Jurkat (ATCC TIB 152), SupT1 (provided by F. Gonzalez-Scarano, Departments of Neurology and Microbiology, University of Pennsylvania, Philadelphia, Pa.), and H9 (Thomas Folk, Retrovirus Disease Branch, Centers for Disease Control and Prevention, Atlanta, Ga.). All T-cell lines were cultured and maintained at 37°C with 5% CO_2 in RPMI 1640 medium. Human neuronal cell lines utilized include IMR-32 (ATCC CCL 127) and SK-N-MC (ATCC HTB 10). Neuronal cell lines were cultured and maintained at 37°C in Eagle minimal essential medium buffered with *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES). Human glial cell lines utilized in these studies include U-373 MG (ATCC HTB 17), Hs683 (ATCC HTB 138), and U-118 (ATCC HTB 15). Glial cell lines were cultured and maintained at 37°C in Eagle minimal essential medium. All medium formulations were supplemented with 10% heat-inactivated fetal calf serum, penicillin (100 U/ml), streptomycin (100 μg/ml), 2 mM L-glutamine, and 0.075% $NaHCO_3$.

Nuclear extract preparation. Nuclear extracts of all cell lines were prepared as described previously (11), with minor modifications. Briefly, cultured cells were collected and nuclei were isolated with hypotonic buffer (10 mM HEPES [pH 7.9, 4°C], 1.5 mM $MgCl_2$, 10 mM KCl, 0.2 mM phenylmethylsulfonyl fluoride, 0.5 mM dithiothreitol) in a Dounce homogenizer (type B) at 4°C. Nuclear proteins were extracted at 4°C with high-salt buffer (20 mM HEPES [pH 7.9, 4°C], 25% glycerol, 1.5 mM $MgCl_2$, 1 M KCl, 0.2 mM EDTA, 0.2 mM phenylmethylsulfonyl fluoride, 0.5 mM dithiothreitol) and dialyzed against 50 volumes of dialysis buffer (20 mM HEPES [pH 7.9, 4°C], 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.2 mM

phenylmethylsulfonyl fluoride, 0.5 mM dithiothreitol) for 5 h at 4°C. Precipitated proteins and debris were removed by centrifugation at 21,000 × *g* for 30 min at 4°C. Protein quantitation of nuclear extracts was performed by the Bio-Rad protein assay (based on the Bradford dye-binding procedure [5]). Extracts were separated into aliquots, frozen in liquid nitrogen, and stored at −80°C at a concentration of 3 mg/ml.

Oligonucleotide synthesis and radiolabeling. Complementary single-stranded oligonucleotides homologous to each of the three HTLV-I 21-bp repeat elements were synthesized (Macromolecular Core Facility, The Pennsylvania State University College of Medicine) and annealed by heating for 10 min at 90°C and cooling to room temperature. For EMS analyses examining the role of flanking DNA in DNA-protein complex formation, double-stranded oligonucleotides containing sequences homologous to the 21-bp repeat elements with native flanking sequences, with restriction endonuclease site-flanking DNA sequences, or without flanking DNA sequences were synthesized. The sequences of the full-length, double-stranded oligonucleotides were as follows: the three 21-bp repeat elements flanked by native HTLV-I LTR sequences, 5′-AGACTAAGGCTCTGACGTCTCCCCCCAGAGG-3′ (I FS), 5′-CAGGCTAGGCCCTGACGTGTCCCCCTGAAGA-3′ (II FS), and 5′-GCCCTCAGGCGTTGACGACAACCCCTCACCT-3′ (III FS); the three 21-bp repeat elements alone with no flanking DNA sequences, 5′-AAGGCTCTGACGTCTCCCCCC-3′ (I), 5′-TAGGCCCTGACGTGTCCCCCT-3′ (II), and 5′-CAGGCGTTGACGACAACCCCT-3′ (III); and the three 21-bp repeat elements flanked by irrelevant DNA sequences in the form of restriction endonuclease sites *Bam*HI and *Hin*dIII, 5′-GATCCAAGGCTCTGACGTCTCCCCCCAAGCT-3′ (I RS), 5′-GATCCTAGGCCCTGACGTGTCCCCCTAAGCT-3′ (II RS), and 5′-GATCCCAGGCGTTGACGACAACCCCTAAGCT-3′ (III RS) (each 21-bp repeat is flanked by the restriction endonuclease sequences for *Bam*HI [5′ end of the 21-bp repeat] and *Hin*dIII [3′ end of the 21-bp repeat]). For mutational analyses of 21-bp repeat III, a synthetic oligonucleotide (21-bp repeat IIIΔII, 5′-GCCCTCAGGCGTTGACGTGTCCCCCTCACCT-3′) in which four nucleotides between conserved domains B and C were changed to the corresponding nucleotides in 21-bp repeat II was synthesized. The 21-bp repeat I corresponds to nucleotides −251 to −231 (promoter distal), 21-bp repeat II corresponds to nucleotides −203 to −183 (central), and 21-bp repeat III corresponds to nucleotides −103 to −83 (promoter proximal) with respect to the transcriptional start site (Fig. 1) (38, 43). Transcription factor binding site oligonucleotides used for competition EMS analyses include (i) a CRE oligonucleotide, 5′-GATTGGCTGACGTCAGAGAGCT-3′, and (ii) an Sp1 binding site oligonucleotide, 5′-GATCGATCGGGGCGGGGCGATC-3′ (Stratagene). The oligonucleotides with 5′ extensions were labeled with [α-^{32}P]dCTP and the Klenow fragment of DNA polymerase I. The blunt-ended double-stranded oligonucleotides were end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase.

Generation of the 223-bp fragment of the HTLV-I LTR encompassing the three 21-bp repeat elements (TRE-1/-2). PCR amplification was utilized to generate a DNA fragment, TRE-1/-2, from the HTLV-I LTR, extending from the *Apa*I restriction endonuclease site at position −268 to the *Nde*I restriction endonuclease site at position −46 relative to the transcriptional start site. The TRE-1/-2 fragment contains the three 21-bp repeat elements and intervening sequences in the context of their native positions within the HTLV-I LTR and comprises sequences previously referred to as the Tax-responsive elements TRE-1 and TRE-2. The two PCR primers were homologous to sequence −268 to −249 (5′ primer) and to sequence −65 to −46 (3′ primer). After PCR amplification, the resulting 223-bp fragment was end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase.

EMS analyses. Standard EMS analyses were performed essentially as described by Garner and Revzin (17). Approximately 50,000 cpm of radiolabeled double-stranded oligonucleotide (0.1 to 1.0 ng) was used in a reaction with 12 μg of protein and 2 μg of poly(dI-dC) in a total reaction volume of 15 μl for 30 min at 30°C. All components were simultaneously mixed by brief centrifugation. For preincubation EMS analyses, the nuclear extracts were incubated with the 2 μg of poly(dI-dC) for 10 min at room temperature prior to incubation with radiolabeled probe DNA. For competition EMS analyses, unlabeled competitor oligonucleotides were simultaneously incubated with nuclear extract, poly(dI-dC), and radiolabeled probe DNA for 30 min at 30°C and subjected to electrophoresis. Following incubation, 2 μl of loading buffer (50% glycerol, 0.1 M EDTA [pH 8.0], 0.1% bromphenyl blue, 0.1% xylene cyanol) was added and each reaction mixture was subjected to electrophoresis in a 4% high-ionic-strength native polyacrylamide gel (prerun for 1.5 h at 100 V) at 30 mA. The polyacrylamide gels were dried under a vacuum at 80°C for 1.5 h prior to autoradiography.

RESULTS

Detection of 21-bp repeat-specific and cell type-specific DNA-protein complexes. Because of the incomplete conservation among the three 21-bp repeat elements, EMS analyses were performed to detect DNA-protein complexes which formed between each of the three 21-bp elements and nuclear factors derived from selected cell lines. Since only 13 of the 21 nucleotides of each 21-bp repeat element are strictly conserved, it is possible that 21-bp repeat-specific DNA-protein complexes form when each of the three 21-bp repeat elements individually react with cellular factors. Standard EMS analyses were performed with nuclear extracts derived from cells of lymphocytic (Jurkat, SupT1, and H9), neuronal (IMR-32 and SK-N-MC), and glial (U-373 MG, Hs683, and U-118) origins and each of the three 21-bp repeat elements present in the U3 region of the HTLV-I LTR. Four low-mobility DNA-protein complexes, designated C1, C2, C3, and C4, were detected when Jurkat-, SupT1-, H9-, IMR-32-, and SK-N-MC-derived nuclear extracts were used in individual reactions with radiolabeled 21-bp repeat I (Fig. 2). In addition, several high-mobility DNA-protein complexes were also detected (Fig. 2, bracket). DNA-protein complexes with similar electrophoretic mobilities were detected in EMS analyses performed with nuclear extracts derived from the glial cell lines U-373 MG (Fig. 2), Hs683 (Fig. 2), and U-118 (data not shown).

Four low-mobility and several high-mobility DNA-protein complexes were also detected when nuclear extracts derived from Jurkat, SupT1, H9, IMR-32, and SK-N-MC cell lines reacted with radiolabeled 21-bp repeat II (Fig. 2). The four low-mobility DNA-protein complexes had electrophoretic mobilities similar to those of the DNA-protein complexes, C1 to C4, detected in EMS analyses utilizing 21-bp repeat I. DNA-protein complexes C1 to C4 were also detected when U-373 MG (Fig. 2), Hs683 (Fig. 2), and U-118 (data not shown) nuclear extracts were incubated with 21-bp repeat II. However, the C1 and C2 DNA-protein complexes detected when either 21-bp repeat I or II was used as a probe in EMS analyses with glial cell-derived nuclear extracts were routinely less abundant. Furthermore, a glial cell type-specific DNA-protein complex with an electrophoretic mobility between those of C2 and C3

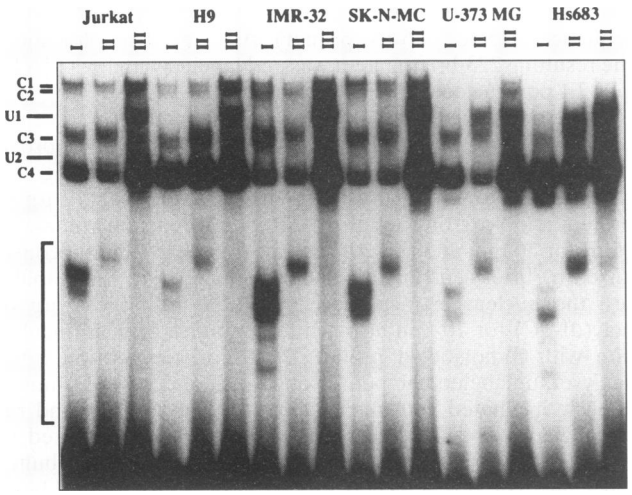

FIG. 2. EMS analyses with nuclear extracts derived from lymphocytic, neuronal, and glial cell lines in reaction with each 21-bp repeat element. Standard EMS reaction mixtures containing 12 μg of lymphocytic (Jurkat or H9), neuronal (IMR-32 or SK-N-MC), or glial (U-373 MG or Hs683) cell-derived nuclear extracts, 2 μg of poly(dI-dC), and 50,000 cpm of radiolabeled 21-bp repeat I, II, or III were incubated at 30°C for 30 min. High-mobility DNA-protein complexes are indicated by the bracket.

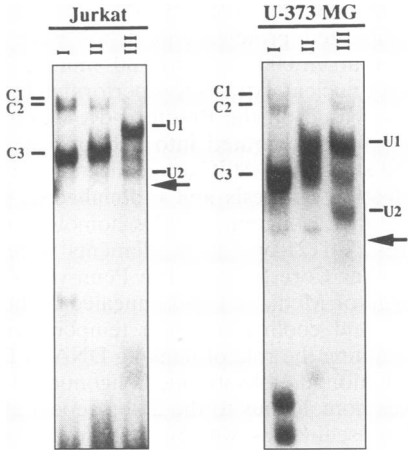

FIG. 3. Preincubation of lymphocytic and glial cell nuclear extracts with nonspecific DNA. Preincubation of EMS reaction mixtures containing 12 μg of Jurkat- or U-373 MG-derived nuclear extract and 2 μg of poly(dI-dC) at room temperature for 10 min was followed by incubation at 30°C for 30 min with 50,000 cpm of radiolabeled 21-bp repeat I, II, or III. The positions of the abrogated C4 complexes detected in EMS reactions with no preincubation are indicated by the arrows. Free radiolabeled DNA probe is not shown.

was detected when nuclear extracts derived from the glial cell lines U-373 MG (Fig. 2), Hs683 (Fig. 2), and U-118 (data not shown) reacted with 21-bp repeat I or II; however, the abundance of the DNA-protein complex was far greater when 21-bp repeat II was utilized as the radiolabeled probe DNA. In addition, several high-mobility DNA-protein complexes, distinct from those detected in assays with radiolabeled 21-bp repeat I, were formed with radiolabeled 21-bp repeat II and each nuclear extract was examined (Fig. 2, bracket).

Collectively, these data demonstrate the formation of both 21-bp repeat I- and II-specific DNA-protein complexes, as well as a glial cell type-specific DNA-protein complex. Since the electrophoretic mobilities of each of the four DNA-protein complexes were similar for each nuclear extract examined, it is likely that the protein constituents involved in the formation of each of the DNA-protein complexes, C1 to C4, are similar regardless of nuclear extract origin. However, subtle differences in the relative abundance and electrophoretic mobility of each DNA-protein complex detected with the series of nuclear extracts examined suggest that although the proteins involved in formation of each DNA-protein complex are likely to be similar for the nuclear extracts examined, they may not necessarily be identical.

Unlike the pattern of four low-mobility DNA-protein complexes detected with 21-bp repeats I and II, at least six DNA-protein complexes were detected when nuclear extracts reacted with radiolabeled 21-bp repeat III. The four low-mobility DNA-protein complexes, C1 to C4, detected when nuclear extracts reacted with 21-bp repeat elements I and II were also detected when Jurkat-, SupT1-, H9-, IMR-32-, SK-N-MC-, U-373 MG-, Hs683-, and U-118-derived nuclear extracts reacted with 21-bp repeat III (Fig. 2; data not shown). In addition to C1 to C4, two unique DNA-protein complexes, designated U1 and U2, were detected when each nuclear extract reacted with 21-bp repeat III (Fig. 2 and 3). The U1 DNA-protein complex (Fig. 2 and 3) had an electrophoretic mobility between those of C2 and C3. Although the electrophoretic mobility of U1 was similar to that of the glial cell-specific DNA-protein complex, competition EMS analyses indicated that the U1 and glial cell-specific DNA-protein complexes comprised distinct protein components (see Fig. 8). The U2 DNA-protein complex was masked by the highly abundant C4 DNA-protein complex but was readily detectable when formation of the C4 DNA-protein complex was abrogated by preincubation of nuclear extracts with nonspecific DNA (Fig. 3). Formation of the U2 DNA-protein complex was not, however, dependent on abrogation of the C4 DNA-protein complex, since the U2 complex could also be detected under standard EMS reaction conditions by extending the electrophoretic separation time (data not shown).

As described above for DNA-protein complexes C1 to C4, each of the unique DNA-protein complexes, U1 and U2, exhibited similar electrophoretic mobilities regardless of nuclear extract origin. Furthermore, the high-mobility DNA-protein complexes detected varied in number and relative abundance contingent on which HTLV-I 21-bp repeat element was utilized as probe DNA. These data indicate that, in addition to 21-bp repeat I- and II-specific and glial cell-specific DNA-protein complexes, there are 21-bp repeat III-specific DNA-protein complexes formed when the 21-bp repeat III element reacts with nuclear extracts derived from selected cell lines. The DNA-protein complexes discussed were routinely detected by utilizing multiple nuclear extract preparations and nuclear extract preparation protocols. In addition, there was no variation in the number or nature of low-mobility DNA-protein complexes formed when poly(dA-dT) or salmon sperm DNA was added as nonspecific DNA in the standard EMS reaction (data not shown).

Specificity of DNA-protein complex formation. Several lines of investigation were examined to determine the specificity of the DNA-protein complexes detected. First, standard EMS analyses were performed, in which increasing nuclear extract protein concentrations were utilized to ensure that U1 and U2 DNA-protein complex formation was detected only with 21-bp repeat III and that glial cell-specific DNA-protein complex

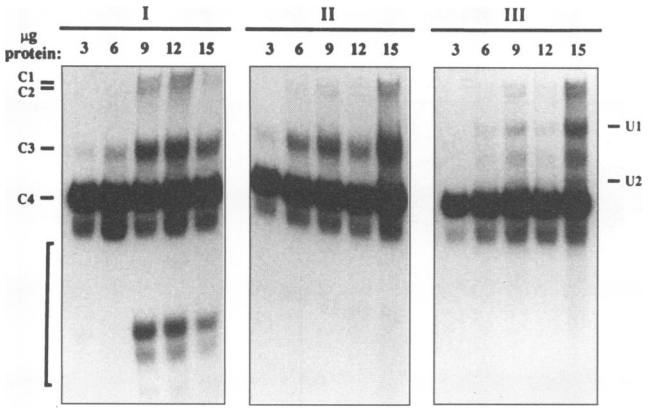

FIG. 4. Effect of protein concentration on DNA-protein complex formation. Standard EMS analyses in which 2 μg of poly(dI-dC), 50,000 cpm of radiolabeled 21-bp repeat I, II, or III, and either 3, 6, 9, 12, or 15 μg of Jurkat (A) or U-373 MG (B) nuclear extract were incubated at 30°C for 30 min. High-mobility DNA-protein complexes are indicated by the bracket.

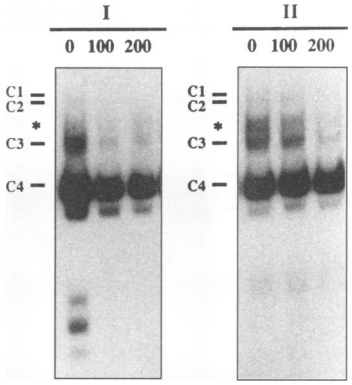

FIG. 5. Cognate competition EMS analyses with glial cell-derived nuclear extracts and 21-bp repeat element I or II. Competition EMS reaction mixtures containing 12 μg of U-373 MG-derived nuclear extract, 2 μg of poly(dI-dC), 50,000 cpm of radiolabeled 21-bp repeat I or II, and either 100- or 200-fold molar excess of unlabeled 21 bp repeat II (as indicated above the autoradiographs) were incubated at 30°C for 30 min. The glial cell-specific DNA-protein complex is indicated by an asterisk.

formation was detected only in assays with glial cell-derived nuclear extracts. As shown in Fig. 4, when as much as 15 μg of Jurkat-derived nuclear extract protein was utilized in standard EMS reactions along with radiolabeled 21-bp repeat I or II, formation of the U1 and U2 DNA-protein complexes was not observed. In fact, amounts of protein up to and including 48 μg of Jurkat-derived nuclear extract did not result in formation of U1 and U2 in reactions with radiolabeled 21-bp repeat element I or II (data not shown). Furthermore, formation of the glial cell-specific DNA-protein complex was not detected by EMS analyses utilizing amounts of protein (Fig. 4) up to and including 48 μg of Jurkat-derived nuclear extract (data not shown). These data demonstrate that formation of the U1, U2, and glial cell-specific DNA-protein complexes was not due to differences in protein availability between EMS reactions or between nuclear extract preparations derived from the indicated cell lines.

In a parallel line of experimentation, increasing levels of radiolabeled probe DNA were utilized in standard EMS reactions to ensure that the inability to detect the U1 and U2 DNA-protein complexes in assays with 21-bp repeats I and II or the glial cell-specific DNA-protein complex in assays with Jurkat-derived nuclear extracts was not due to an inadequate amount of radiolabeled probe DNA. Results from these studies indicated that detection of the 21-bp repeat III-specific complex or glial cell-specific DNA-protein complex was not a result of limited radiolabeled probe levels in the EMS reaction (data not shown). In addition, formation of the 21-bp repeat-specific and glial cell-specific DNA-protein complexes was not dependent on the EMS reaction incubation time, since EMS reactions with extended incubation times (up to and including 60 min) did not result in changes in the number or nature of DNA-protein complexes detected (data not shown). Cumulatively, these results suggest that detection of the 21-bp repeat III-specific and glial cell-specific DNA-protein complexes was in fact due to specific interactions occurring between 21-bp repeat III and cellular factors or 21-bp repeats I and II and glial cell-derived nuclear factors, respectively.

Second, EMS analyses in which the nuclear extracts were preincubated with nonspecific DNA, poly(dI-dC), prior to incubation with each radiolabeled 21-bp repeat element were performed. As shown in Fig. 3, when Jurkat- and U-373 MG-derived nuclear extracts were preincubated with nonspecific DNA prior to incubation with radiolabeled 21-bp repeat I, formation of the C4 DNA-protein complex was abrogated. Similar results were also observed when radiolabeled 21-bp repeat II or III was utilized as probe DNA (Fig. 3). However, the C1 to C3, U1, U2, and glial cell type-specific DNA-protein complexes observed in the standard EMS analyses were detected when the nuclear extracts were preincubated with nonspecific DNA, indicating that these are specific DNA-protein interactions occurring between cellular factors and the 21-bp repeat elements. In addition, and as previously discussed, detection of the U2 DNA-protein complex was evident when formation of the C4 DNA-protein complex was abrogated. Comparable results were also observed when nuclear extracts derived from SupT1, H9, U937, IMR-32, SK-N-MC, Hs683, and U-118 cell lines were utilized in similar experimentation (data not shown). These data suggested that the C4 DNA-protein complex is nonspecific in nature.

A third line of investigation aimed at determining the specificity of DNA-protein complex formation included competition EMS analyses in which cognate DNA sequence was utilized as competitor DNA. As shown in Fig. 5, utilizing U-373 MG-derived nuclear extract and radiolabeled 21-bp repeat I or II, competition with unlabeled cognate sequence resulted in inhibition of C1, C2, C3, and glial cell-specific DNA-protein complex formation. In contrast, formation of the C4 DNA-protein complex was essentially unaffected by the presence of unlabeled competitor DNA. Similar results were obtained from EMS analyses addressing the specificity of U1 and U2 DNA-protein complex formation (data not shown). In competition EMS analyses with Jurkat-derived nuclear extract and radiolabeled 21-bp repeat III, formation of the C1, C2, C3, U1, and U2 DNA-protein complexes was inhibited, whereas formation of C4 was unaffected by the presence of unlabeled cognate competitor DNA. These data further support the probability of the nonspecific nature of the C4 DNA-protein complex and demonstrate that formation of C1, C2, C3, and the glial cell-specific DNA-protein complexes was due to specific interactions between cellular factors and the 21-bp repeat elements.

Dependence of DNA-protein complex formation on flanking

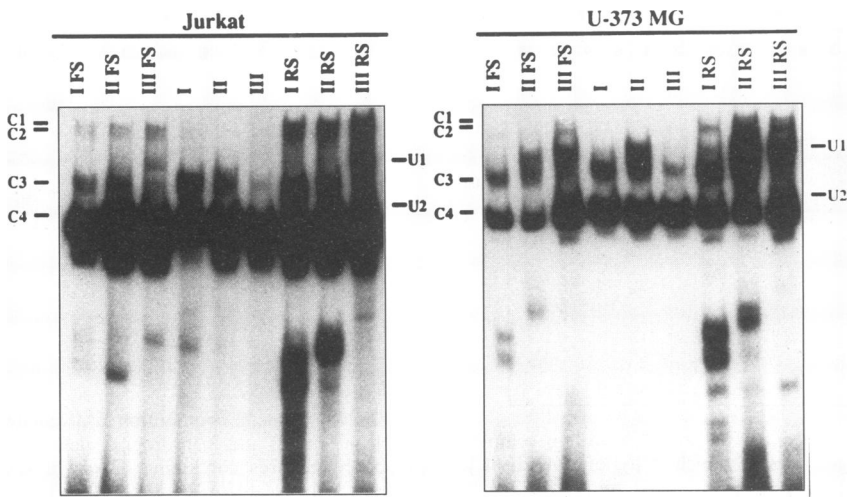

FIG. 6. EMS analyses with lymphocytic and glial cell nuclear extracts and 21-bp repeats with or without flanking DNA. Standard EMS analyses in which 12 μg of Jurkat- or U-373 MG-derived nuclear extract, 2 μg of poly(dI-dC), and 50,000 cpm of radiolabeled 21-bp repeat element native flanking sequences (lanes I FS, II FS, and III FS), no flanking sequences (lanes I, II, and III), or 5′ *Hin*dIII and 3′ *Bam*HI restriction endonuclease site sequences (lanes I RS, II RS, and III RS) were incubated at 30°C for 30 min. Free radiolabeled DNA probe is not shown.

DNA. Since the affinity of DNA-binding proteins for their cognate recognition sites is often affected by the presence of flanking DNA sequences, we have examined the role of flanking DNA in the formation of the DNA-protein complexes detected by the three sets of 21-bp repeat double-stranded oligonucleotides in EMS analyses (see Materials and Methods for the sequences). The first set (I FS, II FS, and III FS) contains sequences corresponding to each of the three 21-bp repeat elements flanked by native HTLV-I LTR sequences, the second set (I, II, and III) corresponds to the 21-bp repeat elements alone with no additional DNA sequences, and the third set (I RS, II RS, and III RS) contains sequences corresponding to each of the three 21-bp repeat elements flanked by the restriction endonuclease sequences for *Bam*HI (5′ end of the 21-bp repeat) and *Hin*dIII (3′ end of the 21-bp repeat). All data presented to this point were generated by utilizing 21-bp repeat elements with native HTLV-I LTR flanking DNA sequences. Consistent with the results presented in Fig. 2, four low-mobility DNA-protein complexes were detected when Jurkat and U-373 MG nuclear extracts reacted with each 21-bp repeat element flanked by native HTLV-I LTR sequences (lanes I FS, II FS, and III FS in Fig. 6). In contrast, incubation of nuclear extracts with the 21-bp repeat elements in the absence of flanking DNA (lanes I, II, and III in Fig. 6) diminished the number of DNA-protein complexes detected, specifically C1, C2, and U1. However, DNA-protein complex formation was restored by 21-bp repeat elements bounded by irrelevant DNA sequences in the form of restriction endonuclease sites *Bam*HI and *Hin*dIII (lanes I RS, II RS, and III RS in Fig. 6). In each case, there was little or no homology between the restriction endonuclease sequences and the native flanking HTLV-I LTR sequences. Therefore, the ability to detect similar DNA-protein complexes regardless of the sequence of the flanking DNA suggests that the requirement for flanking DNA in DNA-protein complex formation is not sequence dependent.

Formation of cell type-dependent DNA-protein complexes with the TRE-1/-2 element. Having demonstrated the formation of at least one glial cell type-specific DNA-protein complex in EMS analyses with HTLV-I 21-bp repeat element I or II and nuclear extracts prepared from U-373 MG (Fig. 2), Hs683 (Fig. 2), or U-118 (data not shown) cells, we proceeded to determine whether glial cell type-specific DNA-protein complexes would form when the 21-bp repeat elements were contained within the context of the HTLV-I LTR. To this end, a DNA probe, the TRE-1/-2 element, corresponding to the sequences contained between the *Apa*I and *Nde*I restriction endonuclease sites within the HTLV-I LTR U3 region (positions −268 to −46) was generated by PCR technology as described in Materials and Methods. The resultant 223-bp PCR product contains the three 21-bp repeat elements and intervening sequences in the context of their native positions within the HTLV-I LTR. When the radiolabeled TRE-1/-2 element was utilized in EMS analyses along with nuclear extracts derived from H9, SupT1, or U-118 cells, several glial cell type-dependent DNA-protein complexes were detected (Fig. 7, asterisks). These data demonstrate that glial cell type-specific DNA-protein complexes could be detected not only with an isolated 21-bp repeat element but also with the intact TRE-1/-2 element.

Specificity of TRE-1/-2 DNA-protein complex formation. The specificity of TRE-1/-2 DNA-protein complex formation was determined by several lines of experimentation. First, the nuclear extracts were preincubated with poly(dI-dC) prior to reaction with radiolabeled TRE-1/-2 probe DNA. Under these conditions, the specific DNA-protein complexes were detected (Fig. 7, thin arrows), while formation of the nonspecific DNA-protein complexes was abrogated (Fig. 7, thick arrows). When increasing concentrations of protein or radiolabeled probe DNA were used in EMS analyses along with lymphocytic or glial cell-derived nuclear extract, the number and nature of DNA-protein complexes detected were unchanged (data not shown). Furthermore, formation of the glial cell-specific DNA-protein complexes was not detected in the titration EMS analyses with lymphocytic nuclear extracts, indicating that formation of these complexes was due to specific interactions between glial cell-derived factors and the TRE-1/-2 element and was not a result of differences in the abundance or availability of protein constituents. Competition EMS analyses were also performed with unlabeled cognate TRE-1/-2 as competitor DNA (Fig. 7). As expected, formation of the nonspecific DNA-protein complexes was unaffected by the

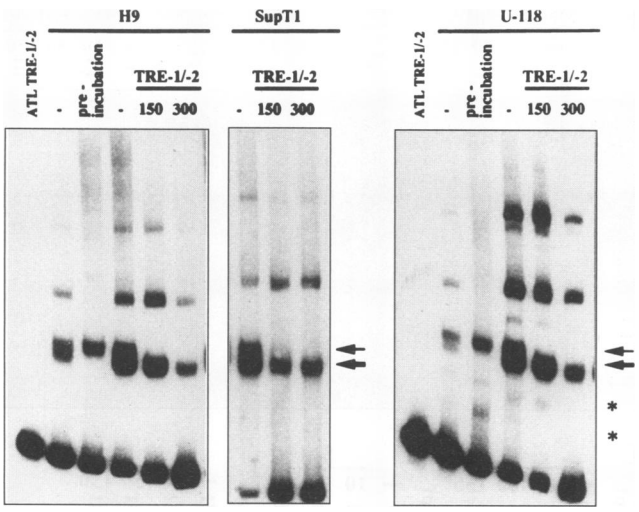

FIG. 7. EMS analyses of the TRE-1/-2 element and nuclear extracts derived from lymphocytic and glial cell-derived nuclear extracts. Standard EMS reaction mixtures containing 3 μg of H9-, SupT1-, or U-118-derived nuclear extract, 2 μg of poly(dI-dC), and 50,000 cpm of radiolabeled TRE-1/-2 element which contains the three 21-bp repeat elements in the context of native HTLV-I LTR sequences were incubated at 30°C for 30 min. Standard EMS reactions performed in the absence of preincubation with competitor DNAs are indicated by dashes. Also shown are cognate EMS analyses in which 150- and 300-fold excesses of unlabeled TRE-1/-2 were used. Thin arrows indicate specific DNA-protein complex formation, thick arrows indicate nonspecific DNA-protein complex formation, and asterisks indicate glial cell-specific DNA-protein complex formation. ATL, adult T-cell leukemia.

presence of cognate competitor DNA, while formation of the specific DNA-protein complexes, including the glial cell-specific DNA-protein complexes, was abrogated.

Characterization of the protein components involved in DNA-protein complex formation with the HTLV-I 21-bp repeat elements. To begin characterization of the protein components involved in DNA-protein complex formation with the 21-bp repeat elements, competition EMS analyses were performed with oligonucleotides containing the binding sites for one of several previously characterized transcription factors, including AP1, AP3, CREB, NF-κB, NF1, and Sp1 (Table 1) (12). Several of the transcription factor binding site consensus sequences exhibit considerable homology to selected regions of the HTLV-I LTR (Table 1) and were therefore considered likely candidates to be involved in DNA-protein complex formation with the HTLV-I 21-bp repeat elements. In addition, previous studies by other investigators (46, 48, 53) have already established that several known transcription factors interact with the 21-bp repeat elements in either the absence or presence of Tax in reactions with nuclear extracts derived from lymphocytes and other cell types of nonneuroglial origin. As shown in Fig. 8, competition EMS analyses were performed with individual radiolabeled 21-bp repeat elements, nuclear extracts derived from Jurkat or U-373 MG cells, and a 10-, 50-, or 150-fold molar excess of unlabeled competitor oligonucleotide. In addition to each 21-bp repeat element being used as radiolabeled probe DNA, radiolabeled transcription factor binding site oligonucleotides were also utilized as probe DNAs in order to determine the electrophoretic mobilities of DNA-protein complexes corresponding to endogenous transcription factors present in the Jurkat or U-373 MG nuclear extracts when they reacted with the cognate binding sites contained within the transcription factor binding site oligonucleotides.

As shown in Fig. 8A, in competition EMS analyses with Jurkat-derived nuclear extracts and radiolabeled 21-bp repeat I element, formation of the C1-C3 DNA-protein complexes was efficiently inhibited by the presence of unlabeled CRE oligonucleotide as competitor DNA. This evidence indicates that a CREB-ATF family member(s), along with individual 21-bp repeat elements, may play a role in the formation of these DNA-protein complexes. Previous studies by other investigators in which multimerized 21-bp repeat elements were utilized as target DNA (34, 35, 48, 53) have also suggested that CREB-ATF family members are involved in specific DNA-protein complex formation. In contrast, formation of the C4 and high-mobility DNA-protein complexes was unaffected by CRE oligonucleotide competitor DNA (Fig. 8A). Similar results were also obtained by utilizing 21-bp repeat II element

TABLE 1. Nucleotide sequences of the 21-bp Tax-responsive elements and of selected transcription factor binding site oligonucleotides

Element	Nucleotide sequence	
	Consensus[a]	Element or oligonucleotide[b]
Tax-responsive elements		▽ ▽
Repeat I		AGACTA TC TCTC CAGAGG
Repeat II		CAGGCTAGGCCCTGACGTGTCCCCCTGAAGA
Repeat III		GCCCTC GT ACAA TCACCT
		Δ Δ′
Transcription factors		
CRE	TGACG(T/C)(C/A)(A/G)	GATTGGCTGACGTCAGAGAGCT
AP1	TGA(G/C)T(C/A)A	CTAGTGATGAGTCAGCCGGATC
AP3	TGTGG(A/T)(A/T)(A/T)	CTAGTGGGACTTTCCACAGATC
NF1	(T/C)GG(C/A)(N_x)GCCAA	ATTTTGGCTTGAAGCCAATATG
NF-κB	GGGA(C/A)TN(T/C)CC	GATCGAGGGGACTTTCCCTAGC
Sp1	(G/T)(G/A)GGC(G/T)(G/A)(G/A)(G/T)	GATCGATCGGGGCGGGGCGATC

[a] Core consensus sequence for selected transcription factor binding sites (12). Alternative nucleotides are indicated as follows: N, any nucleotide; (N_1/N_2), either nucleotide N_1 or nucleotide N_2; or (N_x), a series of nucleotides where x represents the number of nucleotides.

[b] ▽, markers delineating the 5′ and 3′ boundaries of the 21-bp Tax-responsive elements. Nucleotides outside markers represent native HTLV-I LTR flanking sequences. The nucleotide differences among the three 21-bp Tax-responsive elements are indicated, while the nucleotides conserved among the three 21-bp repeats are represented once. Nucleotide sequences of double-stranded oligonucleotides contain a minimal binding site for the corresponding transcription factor. The underlined nucleotides represent the transcription factor binding sites.

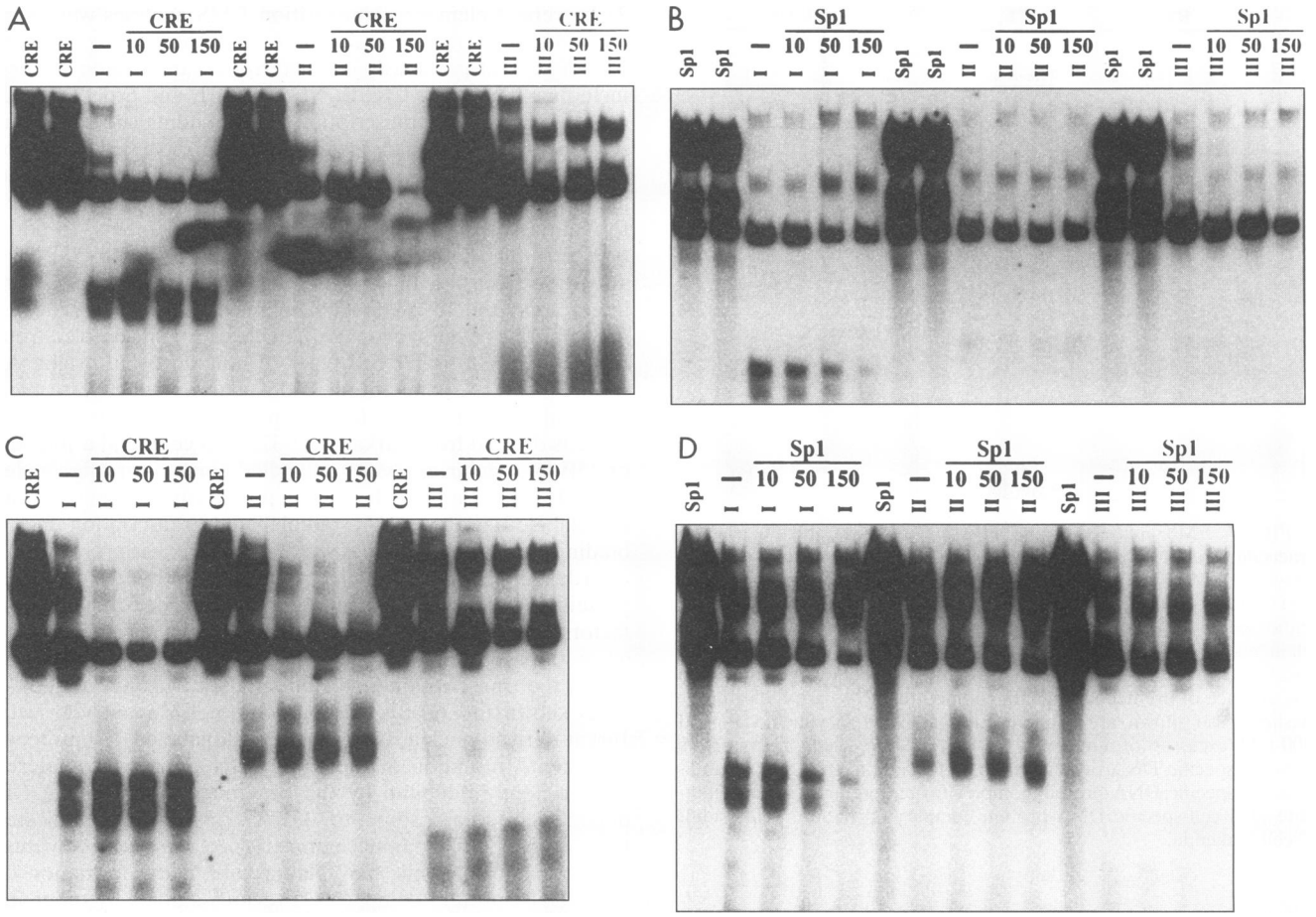

FIG. 8. Competition EMS analyses with unlabeled CRE or Sp1 binding site oligonucleotides and Jurkat- or U-373 MG-derived nuclear extracts. Competition EMS reaction mixtures containing 12 μg of Jurkat-derived (A and C) or U-373 MG-derived (B and D) nuclear extracts, 2 μg of poly(dI-dC), 50,000 cpm of radiolabeled 21-bp repeat I, II, or III, and a 10-, 50-, or 150-fold molar excess of unlabeled CRE (A and B) or Sp1 (C and D) binding site oligonucleotide competitor DNA were incubated at 30°C for 30 min. Shown for comparison are EMS analyses of radiolabeled CRE or Sp1 binding site oligonucleotide in reactions with 12 or 9 μg, respectively, of nuclear extract. Standard EMS reactions performed without preincubation or competitor DNAs are indicated by dashes. Free radiolabeled DNA probe is not shown.

as radiolabeled probe DNA, in that formation of the C1 to C3 DNA-protein complexes was efficiently inhibited with CRE oligonucleotide as the competitor DNA (Fig. 8A), while formation of the C4 and high-mobility DNA-protein complexes was unaffected by the presence of competitor DNA. With radiolabeled CRE oligonucleotide as the probe DNA, DNA-protein complexes which exhibited electrophoretic mobilities similar to those of DNA-protein complexes C1 to C3 with 21-bp repeat elements I and II were detected. This comparative information provides additional evidence suggesting that a member(s) of the CREB-ATF family may participate in formation of these DNA-protein complexes, since the CRE oligonucleotide and the 21-bp repeat oligonucleotides are of similar sizes.

In contrast, when radiolabeled 21-bp repeat III was utilized as probe DNA in similar competition EMS analyses, unlabeled CRE oligonucleotide inhibited only the C1 to C3 DNA-protein complexes efficiently (Fig. 8A). Formation of the U1 and U2 21-bp repeat III-specific DNA-protein complexes was unaffected by CRE oligonucleotide competitor DNA, as was formation of the high-mobility DNA-protein complexes. These results demonstrate that the C1 to C3 DNA-protein complexes detected when 21-bp repeats I, II, and III react with Jurkat-derived nuclear extracts are possibly composed of protein components that recognize the CRE consensus sequence and may include members of the CREB-ATF family of transcription factors. Furthermore, formation of the U1 and U2 21-bp repeat III-specific DNA-protein complexes does not appear to be dependent on a CREB-ATF family member(s).

Competition EMS analyses with nuclear extracts derived from the glial cell line, U-373 MG (Fig. 8B), demonstrated a similar transcription factor binding site oligonucleotide competition profile. When glial cell-derived nuclear extracts reacted with radiolabeled 21-bp repeats I, II, and III in combination with CRE oligonucleotide as the competitor DNA, formation of the C1 to C3 DNA-protein complexes was efficiently inhibited (Fig. 8B). Additionally, formation of the low-mobility, glial cell-specific DNA-protein complex detected in standard EMS reactions with radiolabeled 21-bp repeats I and II was inhibited by CRE oligonucleotide competitor DNA. However, in contrast to the complete abrogation of C1, C2, and C3 DNA-protein complex formation observed in EMS analyses with 10-, 50-, and 150-fold molar excesses of CRE oligonucleotide competitor DNA, comparable levels of CRE competitor resulted in incomplete inhibition of glial cell-specific DNA-protein complex formation.

In competition EMS analyses with 21-bp repeat elements I, II, and III, unlabeled Sp1 binding site oligonucleotide as competitor DNA, and Jurkat-derived nuclear extract (Fig. 8C), the C1 to C4 DNA-protein complexes as well as the high-mobility DNA-protein complexes were unaffected by Sp1 binding site competitor DNA, suggesting that Sp1 is not a component of these DNA-protein complexes. However, when radiolabeled 21-bp repeat III was utilized as probe DNA, formation of the U1 and U2 21-bp repeat III-specific DNA-protein complexes was efficiently inhibited by the presence of Sp1 binding site competitor oligonucleotide, indicating that Sp1 or an Sp1-related factor may play a role in the formation of these DNA-protein complexes. Similar results were obtained with U-373 MG-derived nuclear extract (Fig. 8D); however, competition EMS analyses with U-373 MG-derived nuclear extract, 21-bp repeat III, and competitor Sp1 binding site DNA resulted in only minimal competition of the U1 DNA-protein complex. This result contrasted with those obtained with Jurkat-derived nuclear extracts (Fig. 8C), in which Sp1 binding site competitor DNA efficiently abrogated formation of the U1 DNA-protein complex. These results indicate that while Sp1 or an Sp1-related factor may play a significant role in formation of the U1 and U2 DNA-protein complexes in assays with 21-bp repeat III and lymphocytic nuclear extracts, the role of Sp1 or related factors in DNA-protein complex formation of U1 and U2 with 21-bp repeat III and glial cell-derived nuclear extracts is less apparent. These data further substantiate the existence of cell type-specific differences in DNA-protein complex formation between each of the three 21-bp repeat elements and nuclear extracts derived from selected cell lines.

In similar experimentation performed with Jurkat- or U-373 MG-derived nuclear extracts and AP1, AP3, NF1, and NF-κB binding site oligonucleotides as competitor DNAs (data not shown), DNA-protein complex formation was essentially unaffected, indicating that these transcription factors are not involved in the formation of any of the low- or high-mobility DNA-protein complexes.

Role of nonconserved nucleotides in the formation of 21-bp repeat III-specific and glial cell-specific DNA-protein complexes. The 21-bp enhancer element located within the HTLV-I LTR is imperfectly repeated at three locations within the U3 region. Among the 21-bp repeat elements, there are three strictly conserved domains separated by two nonconserved regions (Fig. 1). To investigate the role of nonconserved nucleotides in the formation of the 21-bp repeat III-specific and glial cell-specific DNA-protein complexes, the 21-bp repeat III element was mutated such that the four nucleotides between conserved domains B and C (ACAA; Fig. 1) were changed to the corresponding sequence of 21-bp repeat II (TGTC, designated 21-bp repeat IIIΔII; see Materials and Methods for the entire sequence). Standard EMS analyses were performed with nuclear extracts derived from Jurkat or U-373 MG cells and radiolabeled 21-bp repeat I, II, III, or IIIΔII. As shown in Fig. 9, when 21-bp repeat IIIΔII was utilized as probe DNA, the U1, 21-bp repeat III-specific complex was not detected, whereas the C1 to C4 DNA-protein complexes were readily detectable. In addition, in EMS analyses with U-373 MG-derived nuclear extracts (Fig. 9), mutation of the sequences between domains B and C in 21-bp repeat IIIΔII resulted in the formation of a DNA-protein complex with an electrophoretic mobility similar to that of the glial cell-specific DNA-protein complex detected by 21-bp repeat elements I and II and glial cell-derived nuclear extracts.

To ensure that the new DNA-protein complex detected by 21-bp repeat IIIΔII and glial cell-derived nuclear extracts was

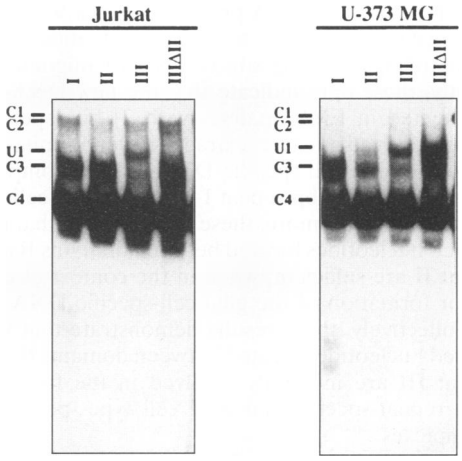

FIG. 9. EMS analyses of 21-bp repeat IIIΔII and nuclear extracts derived from lymphocytic and glial cells. Standard EMS reaction mixtures containing 12 μg of Jurkat- or U-373 MG-derived nuclear extract, 2 μg of poly(dI-dC), and 50,000 cpm of radiolabeled 21-bp repeat I, II, III, or IIIΔII were incubated at 30°C for 30 min.

indeed the glial cell-specific DNA-protein complex and not U1, competition EMS analyses comparable to those previously discussed were performed. Competition EMS analyses with 21-bp repeat IIIΔII, U-373 MG-derived nuclear extract, and CRE or Sp1 binding site competitor oligonucleotide (Fig. 10) demonstrated competition profiles similar to those observed with 21-bp repeat I or II, U-373 MG-derived nuclear extract, and CRE or Sp1 binding site competitor oligonucleotide. That

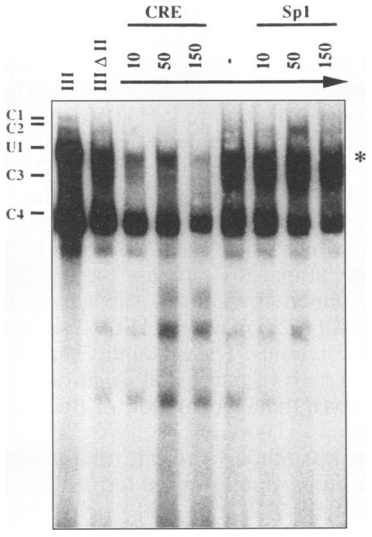

FIG. 10. Competition EMS analyses with unlabeled oligonucleotides containing the binding site for CREB (CRE) or Sp1 and U-373 MG-derived nuclear extracts. Competition EMS reaction mixtures containing 12 μg of U-373 MG-derived nuclear extracts, 2 μg of poly(dI-dC), 50,000 cpm of radiolabeled 21-bp repeat III or IIIΔII, and a 10-, 50-, or 150-fold molar excess (marked above lanes) of unlabeled CRE or Sp1 binding site oligonucleotide competitor DNA were incubated at 30°C for 30 min. A standard EMS reaction performed without preincubation or competitor DNAs is indicated by the dash. Free radiolabeled DNA probe is not shown. *, glial cell-specific DNA-protein complex.

is, formation of the new DNA-protein complex was inhibited by the presence of CRE competitor oligonucleotide but not by the presence of Sp1 binding site competitor oligonucleotide. Consequently, these data indicate that the new DNA-protein complex detected in EMS analyses with 21-bp repeat IIIΔII and glial cell-derived nuclear extracts is very similar, if not identical, to the glial cell-specific DNA-protein complex previously detected with 21-bp repeat I or II and glial cell-derived nuclear extracts. Furthermore, these data indicate that the four nonconserved nucleotides located between domains B and C in 21-bp repeat II are sufficient, when in the context of domains B and C, for formation of the glial cell-specific DNA-protein complex. Collectively, these results demonstrate that the four nonconserved nucleotides located between domains B and C in 21-bp repeat III are integrally involved in the formation of both 21-bp repeat-specific and glial cell type-specific DNA-protein complexes.

DISCUSSION

A variety of cellular proteins has been shown to interact with the HTLV-I LTR, specifically with the 21-bp Tax-responsive elements (33, 46, 48, 53). Since the interaction of cellular factors with each of the 21-bp repeats is critical to both basal and Tax-mediated HTLV-I LTR-directed expression, identification and characterization of these DNA-protein interactions are critical to determining the precise components involved in HTLV-I LTR-directed expression in immune and nervous system cell types. Furthermore, the analysis of cell type-dependent interactions will contribute to dissecting the complex pathway leading to HTLV-I cell type-specific expression, an event which may directly impact on the regulation of (i) productive viral replication; (ii) establishment, maintenance, and reactivation of latent viral infection; and (iii) viral pathogenicity within targeted cell populations.

Although considerable effort has been exerted to identify and characterize the cellular factors of lymphocytic origin which interact with the HTLV-I regulatory unit, neuroglial cell-derived proteins which interact with HTLV-I LTR sequences have not been reported. HTLV-I has been shown in vitro to infect cells of nonlymphoid origin, including a number of cell types of nervous system origin (1, 13, 22–24, 28, 32, 49, 52). Furthermore, PCR amplification studies have demonstrated the presence of HTLV-I DNA sequences in several regions of the brain and spinal cord of HTLV-I-infected tropical spastic paraparesis patients (4, 25, 27), contrasting with results obtained for control subjects and adult T-cell leukemia patients (4, 27) in which HTLV-I sequences were not detected in these regions. The possibility that HTLV-I infects and replicates in cells of neuroglial origin warrants investigation of the transcriptional regulation of the HTLV-I LTR in these cell types.

Results from our studies indicate that similar as well as unique DNA-protein complexes are formed by cellular factors of lymphocytic, neuronal, and glial origins and each of the three 21-bp repeat elements. All of the C1, C2, and C3 DNA-protein complexes detected from nuclear extracts in reactions with individual 21-bp repeat elements exhibited similar electrophoretic mobilities, suggesting that at least some of the DNA-protein complexes formed between nuclear factors and each of the 21-bp repeat elements are composed of similar but not necessarily identical proteins. In contrast, at least two of the DNA-protein complexes detected, U1 and U2, were 21-bp repeat III specific, indicating that each of the 21-bp repeat elements is unique with respect to its ability to interact with select cellular factors.

In addition to the 21-bp repeat-specific nature of certain DNA-protein complexes detected, at least one DNA-protein complex which was detected in EMS analyses with 21-bp repeats I and II and glial cell-derived nuclear extracts was glial cell type specific. Formation of glial cell-specific DNA-protein complexes was also detected by a fragment of the HTLV-I LTR containing the three 21-bp repeat elements in the context of native HTLV-I LTR sequences (TRE-1/-2), demonstrating that glial cell-specific DNA-protein complexes form with not only an isolated 21-bp repeat element but also with the intact TRE-1/-2 element. Site-directed mutagenesis studies are under way to determine the specific sequences within TRE-1/-2 which are involved in glial cell-specific DNA-protein complex formation, as are studies to identify the protein components of these complexes. Several possibilities exist for glial cell-derived proteins to play a role in glial cell-specific HTLV-I LTR-directed expression via interaction with the 21-bp repeat elements. For example, formation of cell type-specific DNA-protein complexes with the 21-bp repeat elements may influence basal HTLV-I LTR-directed transcription, thereby enhancing or under certain circumstances suppressing the initial phase of viral replication. Furthermore, the preferential interaction of cell type-specific factors with the 21-bp repeat elements may affect the overall level of Tax-mediated *trans* activation within select target cell populations.

Studies aimed at examining the role of flanking DNA in DNA-protein complex formation demonstrated that the 21-bp repeat alone was sufficient for formation of the glial cell-specific DNA-protein complex. Formation of the 21-bp repeat-specific complexes, in contrast, was dependent on the presence of flanking DNA; however, this requirement was not sequence specific. Mutational analyses of 21-bp repeat III indicated that the four nonconserved nucleotides between domains B and C (Fig. 1) were necessary for formation of both 21-bp repeat III-specific and glial cell-specific DNA-protein complexes. Replacement of these nucleotides in 21-bp repeat III (ACAA) with those of 21-bp repeat II (TGTC) abrogated U1 DNA-protein complex formation but reconstituted glial cell-specific complex formation, indicating that TGTC in the context of domains B and C is sufficient for glial cell-specific DNA-protein complex formation.

Further evidence supporting the unique nature of each individual 21-bp repeat element and the presence of glial cell-specific complex formation was obtained through competition EMS analyses with a series of unlabeled transcription factor binding site oligonucleotides. Initial studies were directed at determining the role of these factors in DNA-protein complex formation with the 21-bp repeat elements, since each 21-bp repeat contains strictly conserved core sequences (Fig. 1) that exhibit homology to the CREB, AP1, NF-κB, and Sp1 binding site consensus sequences. On the basis of these studies, a member(s) of the CREB-ATF family of cellular transcription factors was determined to be critical for formation of C1, C2, and C3, as well as for formation of the glial cell-specific complex detected in assays using 21-bp repeat I or II and glial cell-derived nuclear extracts. A higher level of CRE competitor was required to interfere with glial cell-specific complex formation than to completely abrogate formation of C1, C2, and C3, indicating that the CREB-ATF-related factor involved in glial cell-specific complex formation is not identical to that participating in formation of C1, C2, and C3.

The CREB-ATF family consists of a number of structurally related transcription factors which bind to similar DNA elements; however, individual members show distinct transcriptional effector functions (6, 21). Additionally, cell type-specific isoforms of CREB-ATF family members exist (30), supporting

the hypothesis that a glial cell-specific CREB-ATF-related factor(s) is involved in glial cell-specific complex formation. This proposal may account for the observed differences in the CRE competition profiles noted for C1, C2, C3 and the glial cell-specific DNA-protein complex. Alternatively, another CREB-related factor, the CRE modulator (CREM) protein, may participate either alone or in conjunction with CREB-ATF factors in formation of the glial cell-specific complex. Like CREB, several CREM isoforms which are capable of interacting with CREB-ATF factors and binding to the CRE motif, thereby enhancing or blocking transcriptional activation mediated by CREB factors, are produced (10, 15, 29). Furthermore, certain CREM isoforms exhibit tissue type-specific expression and are developmentally regulated (10, 15, 29). Since one of the primary sites of CREM expression includes the central nervous system (10, 15, 29), current research is focused on determining the role of CREM isoforms in formation of the glial cell-specific DNA-protein complex as well as the C1, C2, and C3 complexes.

In addition to the involvement of a CREB-ATF family member(s) in DNA-protein complex formation with the 21-bp repeat elements, competition EMS analyses indicated that the transcription factor Sp1 or an Sp1-like factor participates in U1 and U2 formation. However, a significant difference was observed in the abilities of Sp1 binding site oligonucleotide to inhibit formation of U1 and U2 in competition EMS analyses performed with Jurkat- and U-373 MG-derived nuclear extracts. The presence of Sp1 binding site oligonucleotide completely abrogated formation of U1 and U2 in analyses with Jurkat-derived nuclear extracts, whereas similar assay conditions resulted in incomplete inhibition with U-373 MG-derived nuclear extracts. Despite the presence of Sp1 in most cell types, tissue-specific expression of Sp1 as well as tissue-specific enhancement of Sp1-mediated gene expression have been reported previously (2, 26). Cell- and/or tissue-specific expression of Sp1 or an Sp1-related factor may account for the observed differences in Sp1 binding site oligonucleotide competition profiles in assays with 21-bp repeat III and nuclear extracts of lymphocytic or glial origin. Alternatively, the observed U1 complex detected in assays with U-373 MG-derived nuclear extracts could be two or more comigrating complexes, only one of which is dependent on Sp1 for formation.

In summary, we present evidence that 21-bp repeat-specific and glial cell-specific DNA-protein complexes can form between the HTLV-I 21-bp Tax-responsive elements and cellular factors of immune and nervous system origins. Current studies are directed at identifying and characterizing the factors involved in formation of the glial cell-specific complex, assessing the functional role of the 21-bp repeats in cells of lymphocytic and glial cell origins, and defining the functional role of neuroglial cell-specific protein factors in cell type-specific expression.

ACKNOWLEDGMENTS

We gratefully acknowledge the technical assistance of Wang Dian.

This research was supported by a grant awarded to B.W. from the National Cancer Institute (Public Health Service grant CA 54559).

REFERENCES

1. **Akagi, T., Y. Hoshida, T. Yoshino, N. Teramoto, E. Kondo, K. Hayashi, and K. Takahashi.** 1992. Infectivity of human T-cell leukemia virus type I to human nervous tissue cells *in vitro*. Acta Neuropathol. **84:**147–152.
2. **Bessereau, J., D. Mendelzon, C. LePoupon, M. Fiszman, J. Changeux, and J. Piette.** 1993. Muscle-specific expression of the acetylcholine receptor alpha-subunit gene requires both positive and negative interactions between myogenic factors, Sp1 and GBF factors. EMBO J. **12:**443–449.
3. **Bhagavati, S., G. Ehrlich, R. W. Kula, S. Kwok, J. Sninsky, V. Udani, and B. J. Poiesz.** 1988. Detection of human T-cell lymphoma/leukemia virus type I DNA and antigen in spinal fluid and blood of patients with chronic progressive myelopathy. N. Engl. J. Med. **318:**1141–1147.
4. **Bhigjee, A. I., C. A. Wiley, W. Wachsman, T. Amenomori, D. Pirie, P. L. A. Bill, and I. Windsor.** 1991. HTLV-I-associated myelopathy: clinicopathologic correlation with localization of provirus to the spinal cord. Neurology **41:**1990–1992.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
6. **Brindle, P. K., and M. R. Montminy.** 1992. The CREB family of transcription activators. Curr. Opin. Genet. Dev. **2:**199–204.
7. **Celander, D., and W. A. Haseltine.** 1984. Tissue-specific transcription preference as a determinant of cell tropism and leukaemogenic potential of murine retroviruses. Nature (London) **312:**159–162.
8. **Chatis, P. A., C. A. Holland, J. W. Hartley, W. P. Rowe, and N. Hopkins.** 1983. Role for the 3′ end of the genome in determining disease specificity of Friend and Moloney murine leukemia viruses. Proc. Natl. Acad. Sci. USA **80:**4408–4411.
9. **Corboy, J. R., J. M. Buzy, M. C. Zink, and J. E. Clements.** 1992. Expression directed from HIV long terminal repeats in the central nervous system of transgenic mice. Science **258:**1804–1808.
10. **Delmas, V., B. M. Laoide, D. Masquilier, R. P. de Groot, N. S. Foulkes, and P. Sassone-Corsi.** 1992. Alternative usage of initiation codons in mRNA encoding the cAMP-responsive-element modulator generates regulators with opposite functions. Proc. Natl. Acad. Sci. USA **89:**4226–4230.
11. **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.
12. **Faisst, S., and S. Meyer.** 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res. **20:**3–26.
13. **Fanauchi, M., T. Saida, K. Saida, M. Makajima, S. Matsuda, E. Nishiguchi, and H. Nishitani.** 1989. Cytopathic effects of HTLV-I infection in human neural cell lines, p. 333–336. *In* G. C. Roman, J.-C. Vernant, and M. Osame (ed.), HTLV-I in the nervous system. Alan R. Liss, Inc., New York.
14. **Felber, B. K., H. Paskalis, C. Kleinman-Ewing, F. Wong-Staal, and G. N. Pavlakis.** 1985. The pX protein of HTLV-I is a transcriptional activator of its long terminal repeats. Science **229:**675–679.
15. **Foulkes, N. S., E. Borelli, and P. Sassone-Corsi.** 1991. CREM gene: use of alternative DNA-binding domains generates multiple antagonists of cAMP-induced transcription. Cell **64:**739–749.
16. **Fujisawa, J.-I., M. Toita, T. Yoshimura, and M. Yoshida.** 1991. The indirect association of human T-cell leukemia virus *tax* protein with DNA results in transcriptional activation. J. Virol. **65:**4525–4528.
17. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. Nucleic Acids Res. **9:**3047–3060.
18. **Gessain, A., J. C. Vernant, L. Maurs, F. Barin, O. Gout, A. Celander, and G. de The.** 1985. Antibodies to human T-lymphotropic virus type I in patients with tropical spastic paraparesis. Lancet **ii:**407–410.
19. **Giam, C.-Z., and Y.-L. Xu.** 1989. HTLV-I *tax* gene product activates transcription via pre-existing cellular factors and cAMP responsive element. J. Biol. Chem. **264:**15236–15241.
20. **Gonzalez-Dunia, D., G. Grimber, P. Briand, M. Brahic, and S. Ozden.** 1992. Tissue expression pattern directed in transgenic mice by the LTR of an HTLV-I provirus isolated from a case of tropical spastic paraparesis. Virology **187:**705–710.
21. **Hai, T.-Y., F. Liu, W. Coukos, and M. Green.** 1989. Transcription factor ATF cDNA clones: an extensive family of leucine zipper proteins able to selectively form DNA binding heterodimers. Genes Dev. **3:**2083–2090.

22. **Hirayama, M., T. Miyadai, T. Yokochi, K. Sato, T. Kubota, M. Iida, and N. Fujiki.** 1988. Infection of human T-lymphotropic virus type I to astrocytes in vitro with induction of the class II major histocompatibility complex. Neurosci. Lett. **92:**34–39.
23. **Hoffman, P. M., S. Dhib-Jalbut, J. A. Mikovits, D. S. Robbins, A. L. Wolf, G. K. Bergey, N. C. Lohrey, O. S. Weislow, and F. W. Ruscetti.** 1992. Human T-cell leukemia virus type I infection of monocytes and microglial cells in primary human cultures. Proc. Natl. Acad. Sci. USA **89:**11784–11788.
24. **Hoxie, J. A., D. M. Matthews, and D. B. Cines.** 1984. Infection of human endothelial cells by human T-cell leukemia virus type I. Proc. Natl. Acad. Sci. USA **81:**7591–7595.
25. **Iannone, R., M. P. Sherman, P. E. B. Rodgers-Johnson, M. A. Beilke, C. A. Mora, R. M. Amin, S. R. Tinsley, L. D. Papsidero, B. J. Poiesz, and C. J. Gibbs.** 1992. HTLV-I DNA sequences in CNS tissue of a patient with tropical spastic paraparesis and HTLV-I myelopathy. J. Acquired Immune Defic. Syndr. **5:**810–816.
26. **Kingsley, C., and A. Winoto.** 1992. Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. Mol. Cell. Biol. **12:**4251–4261.
27. **Kira, J.-I., Y. Itoyama, Y. Koyanagi, J. Tateishi, M. Kiskikawa, S.-I. Akizuki, I. Kobayashi, N. Toki, K. Sueishi, H. Sato, Y. Sakaki, N. Yamamoto, and I. Goto.** 1992. Presence of HTLV-I proviral DNA in central nervous system of patients with HTLV-I-associated myelopathy. Ann. Neurol. **31:**39–45.
28. **Koralnik, I. J., J. F. Lemp, R. C. Gallo, and G. Franchini.** 1992. In vitro infection of human macrophages by human T-cell-leukemia/lymphotropic virus type I (HTLV-I). AIDS Res. Hum. Retroviruses **8:**1845–1849.
29. **Laoide, B. M., N. S. Foulkes, F. Schlotter, and P. Sassone-Corsi.** 1993. The functional versatility of CREM is determined by its modular structure. EMBO J. **12:**1179–1191.
30. **Lee, K.** 1992. Dimeric transcription factor families: it takes two to tango but who decides on partners and the venue? J. Cell Sci. **103:**9–14.
31. **Li, Y., E. Golemis, J. W. Hartley, and N. Hopkins.** 1987. Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides. J. Virol. **61:**693–700.
32. **Macchi, B., B. Caronti, M. Pezzella, E. Bonmassar, and G. Lauro.** 1991. Effect of human T lymphotropic retrovirus-I exposure on cultured human glioma cell lines. Acta Neuropathol. **81:**670–674.
33. **Matthews, M.-A. H., R.-B. Markowitz, and W. S. Dynan.** 1992. In vitro activation of transcription by the human T-cell leukemia virus type I *tax* protein. Mol. Cell. Biol. **12:**1986–1996.
34. **Nakamura, M., M. Niki, K. Ohtani, and K. Sugamura.** 1989. Differential activation of the 21-base-pair enhancer element of human T-cell leukemia virus type I by its own trans-activator and cyclic AMP. Nucleic Acids Res. **17:**5208–5221.
35. **Niki, M., K. Ohtani, M. Nakamura, and K. Sugamura.** 1992. Multistep regulation of enhancer activity of the 21-base-pair element of human T-cell leukemia virus type I. J. Biol. Chem. **265:**2537–2542.
36. **Nyborg, J. K., and W. S. Dynan.** 1990. Interaction of cellular proteins with the human T-cell leukemia virus type I transcriptional control region, part I. J. Biol. Chem. **265:**8230–8236.
37. **Paquette, Y., Z. Hanna, P. Savard, R. Brousseau, Y. Robitaille, and P. Jolicoeur.** 1989. Retrovirus-induced murine motor neuron disease: mapping the determinant of spongiform degeneration within the envelope gene. Proc. Natl. Acad. Sci. USA **86:**3896–3900.
38. **Park, R. E., W. A. Haseltine, and C. A. Rosen.** 1988. A nuclear factor is required for transactivation of HTLV-I gene expression. Oncogene **3:**275–279.
39. **Paskalis, H., B. K. Felber, and G. N. Pavlakis.** 1986. *cis*-acting sequences responsible for the transcriptional activation of human T-cell leukemia virus type I constitute a conditional enhancer. Proc. Natl. Acad. Sci. USA **83:**6558–6562.
40. **Popovic, M., P. S. Sarin, M. Robert-Gurroff, V. S. Kalyanaraman, D. Mann, J. Minowada, and R. C. Gallo.** 1983. Isolation and transmission of human retrovirus (human T-cell leukemia virus). Science **219:**856–859.
41. **Rosen, C. A., W. A. Haseltine, J. Lenz, R. Ruprecht, and M. W. Cloyd.** 1985. Tissue selectivity of murine leukemia virus infection is determined by long terminal repeat sequences. J. Virol. **55:**862–866.
42. **Rosen, C. A., R. Park, J. G. Sodroski, and W. A. Haseltine.** 1987. Multiple sequence elements are required for regulation of human T-cell leukemia virus gene expression. Proc. Natl. Acad. Sci. USA **84:**4919–4923.
43. **Rosen, C. A., J. G. Sodroski, and W. A. Haseltine.** 1985. Location of *cis*-acting regulatory sequences in the human T-cell leukemia virus type I long terminal repeat. Proc. Natl. Acad. Sci. USA **82:**6502–6506.
44. **Seiki, M., J.-I. Inoue, T. Takeda, and M. Yoshida.** 1986. Direct evidence that $p40^x$ of human T-cell leukemia virus type I is a *trans*-acting transcriptional activator. EMBO J. **5:**561–565.
45. **Shimotohno, K., M. Takano, T. Teruuchi, and M. Miwa.** 1986. Requirement of multiple copies of a 21-nucleotide sequence in the U3 regions of human T-cell leukemia virus type I and type II long terminal repeats for trans-acting activation of transcription. Proc. Natl. Acad. Sci. USA **83:**8112–8116.
46. **Sodroski, J. G.** 1992. The human T-cell leukemia virus (HTLV) transactivator (Tax) protein. Biochim. Biophys. Acta **1114:**19–29.
47. **Sodroski, J. G., C. A. Rosen, and W. A. Haseltine.** 1984. *trans*-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. Science **225:**381–385.
48. **Suzuki, T., J.-I. Fujisawa, M. Toita, and M. Yoshida.** 1993. The trans-activator *Tax* of human T-cell leukemia virus type 1 (HTLV-1) interacts with cAMP-responsive element (CRE) binding and CRE modulator proteins that bind to the 21-base-pair enhancer of HTLV-1. Proc. Natl. Acad. Sci. USA **90:**610–614.
49. **Watabe, K., T. Saida, and S. U. Kim.** 1989. Human and simian glial cells infected by human T-lymphotropic virus type I in culture. J. Neurpathol. Exp. Neurol. **48:**610–619.
50. **Watanabe, T., M. Seiki, and M. Yoshida.** 1984. HTLV type I (U.S. isolate) and ATLV (Japanese isolate) are the same species of human retrovirus. Virology **133:**238–241.
51. **Wong-Staal, F., and R. C. Gallo.** 1985. The family of human T-lymphotropic leukemia viruses: HTLV-I as a cause of adult T cell leukemia and HTLV-III as the cause of acquired immunodeficiency syndrome. Blood **65:**253–263.
52. **Yamada, M., K. Watabe, T. Saida, and S. U. Kim.** 1991. Increased susceptibility of human fetal astrocytes to human T-lymphotropic virus type I in culture. J. Neuropathol. Exp. Neurol. **50:**97–107.
53. **Zhao, L. J., and C.-Z. Giam.** 1991. Human T-cell lymphotropic virus type I (HTLV-I) transcriptional activator, *Tax*, enhances CREB binding to the HTLV-I 21-base-pair repeats by protein-protein interaction. Proc. Natl. Acad. Sci. USA **89:**7070–7074.