Journal of NeuroVirology (1995) 1, 62–77

Characterization of a glial cell-specific DNA-protein complex formed with the human T cell lymphotropic virus type I (HTLV-I) enhancer

Renee Wessner, Maribeth Tillmann-Bogush and Brian Wigdahl

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

Human T cell lymphotropic virus type I (HTLV-I) encodes the *trans*-activator, Tax, which facilitates viral transcription from three 21 bp repeated elements within the U3 region of the long terminal repeat (LTR). Electrophoretic mobility shift (EMS) analyses utilizing double-stranded (ds) oligonucleotides homologous to each of the 21 bp repeats and nuclear extracts derived from selected cell lines of lymphocytic, neuronal, and glial origin have demonstrated the differential binding of cellular factors to each of the three 21 bp repeats. Specifically, both a glial cell-specific DNA-protein complex (designated GCS) and 21 bp repeat-specific DNA-protein complexes (designated U1 and U2) were detected. The formation of the GCS DNA-protein complex may involve activating transcription factor (ATF)/cAMP-response element (CRE) binding protein (CREB) family member(s) while the formation of the U1 and U2 DNA-protein complexes may involve an Sp1-related factor. In addition, three ATF-CREB-related DNA-protein complexes common to each individual 21 bp repeat (designated C1–C3) were also detected. However, we demonstrated that the abundance of the C1 and C2 DNA-protein complexes detected with the individual 21 bp repeats and glial cell nuclear extract was relatively low compared to that obtained with lymphocyte, monocyte, or neuronal nuclear extracts. We also have demonstrated that the ATF-CREB factors participating in formation of the GCS DNA-protein complex are distinct from those participating in formation of the C1–C3 DNA-protein complexes. Based on nucleotide sequence requirements and immunoreactivity, we suggest that the GCS DNA-protein complex may contain a novel glial cell type specific ATF-CREB-related factor(s). Furthermore, we demonstrate that the CRE modulator (CREM) protein in conjunction with the ATF/CREB factor, CREBP1, interact with each of the three 21 bp repeats to form the C3 DNA-protein complex. However, the abundance of the C3 DNA-protein complex formed utilizing the promoter proximal repeat is dramatically lower compared to either of the other two 21 bp repeat elements. Based on these observations, we suggest that the differential binding of cellular factors to each of the three 21 bp repeat elements may play a role in basal as well as Tax-mediated LTR-directed transcription within cell populations of either immune or nervous system origin.

Keywords: HTLV-I; Tax; retroviruses; neuropathogenesis; transcription; glial cells

Introduction

Human T cell lymphotropic virus type I (HTLV-I) has been determined to be an etiologic agent of both adult T cell leukemia (ATL) and a slowly progressive neurologic disorder, tropical spastic paraparesis (TSP) (Popovic *et al*, 1983; Watanabe *et al*, 1984; Gessain *et al*, 1985; Wong-Staal and Gallo, 1985; Bhagavati *et al*, 1988; Bhigjee *et al*, 1991; Akagi *et al*, 1992; Iannone *et al*, 1992; Kira *et al*, 1992; Sodroski, 1992). However, the exact nature of the viral and host factors involved in determining the outcome of HTLV-I infection (malignancy versus neurologic dysfunction) remains unresolved. Any aspect of the retroviral life cycle including entry, reverse transcription, integration, transcription, and

Correspondence: B Wigdahl
Received 20 October 1994; revised 23 November 1994; accepted 23 November 1994

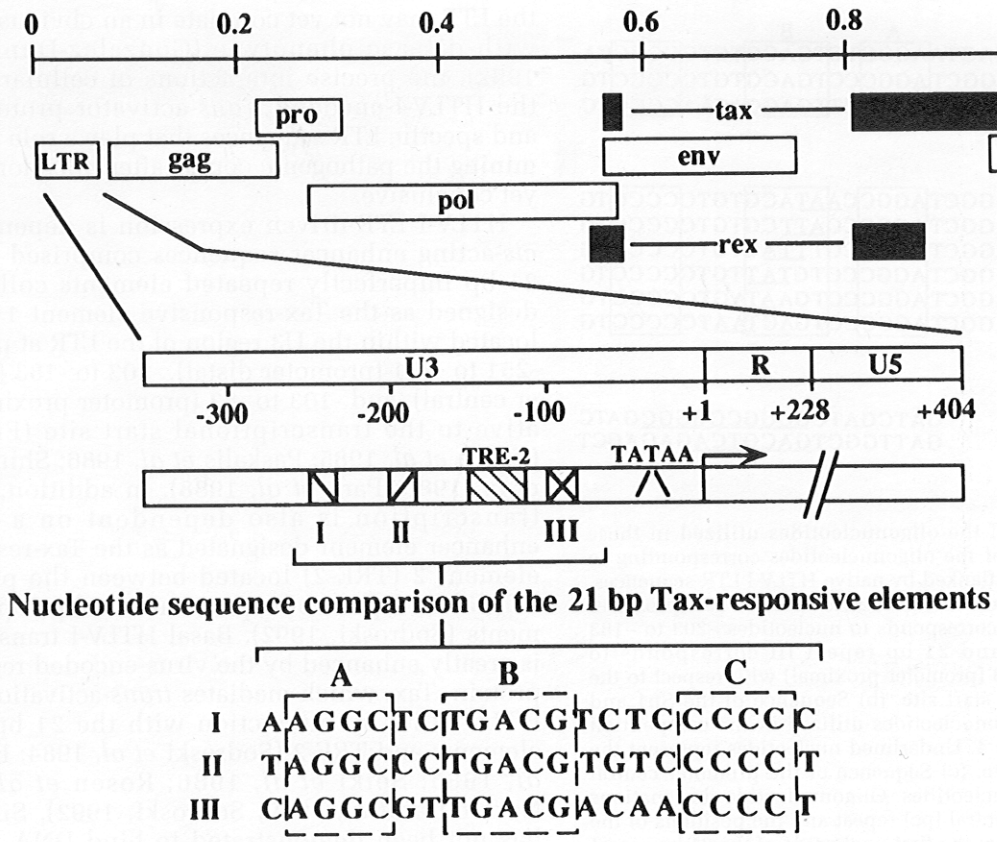

Figure 1 Genomic and HTLV-I LTR structure. The HTLV-I genomic structure and an expanded view of the LTR region is illustrated. The Tax-responsive 21 bp repeats (collectively designated as the Tax-responsive element 1 or TRE-1) are located within the U3 region of the LTR at positions – 251 to –231 (promoter distal), –203 to –183 (promoter central), and –103 to –83 (promoter proximal) relative to the transcriptional start site. Also shown is the Tax-responsive element 2 designated TRE-2 located between the promoter central and promoter proximal 21 bp repeats located at positions – 117 to –163 relative to the transcriptional start site. A nucleotide sequence comparison of the three 21 bp repeats is also shown.

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 retroviral-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. While a wealth of information exists concerning the regulation of HTLV-I gene expression in cells of lymphoid origin, very little is known regarding the neuroglial factors that may interact with the viral regulatory element, designated the long terminal repeat (LTR), and consequently modulate viral gene expression within the central nervous system (CNS).

As with other retroviruses, the integrated HTLV-I genome is flanked by non-coding LTR sequences comprised of three regions, U3, R, and U5 (Figure 1), which contain the information essential for the regulation of integration, transcription, and replication. Evidence suggests 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 infection (Chatis *et al*, 1983; Celander and Haseltine, 1984; Li *et al*, 1987; Rosen *et al*, 1985). For example, it has been suggested that the *in vitro* host range of selected murine leukemia viruses (MuLVs) is determined by the LTR, specifically the U3 region (Chatis *et al*, 1983; Celander and Haseltine, 1984; Rosen *et al*, 1985; Li *et al*, 1987). 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, the pathogenicity (Celander and Haseltine, 1984; Rosen *et al*, 1985; Li *et al*, 1987; Paquette *et al*, 1989). In addition, the promoters of several retroviruses, including Moloney and Friend MuLVs and human immunodeficiency virus (HIV), have been implicated in cell-type specific expression (Chatis *et al*, 1983; Celander and Haseltine, 1984; Rosen *et al*, 1985; Li *et al*, 1987;

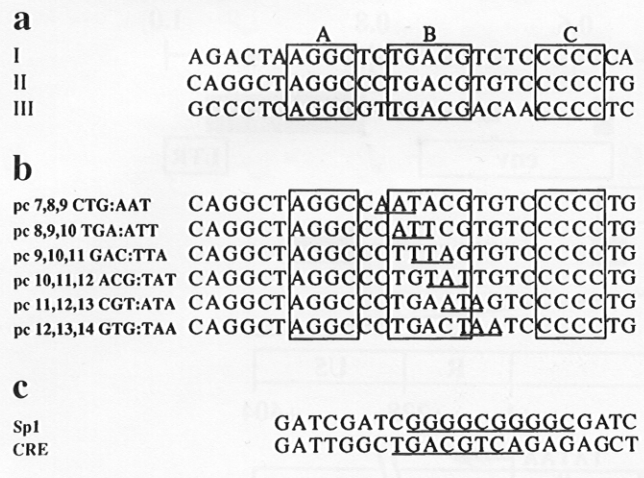

Figure 2 Sequence of the oligonucleotides utilized in these studies. (a) Sequence of the oligonucleotides corresponding to the three 21 bp repeats flanked by native HTLV-I LTR sequences. 21 bp repeat I corresponds to nucleotides –251 to –231 (promoter distal), 21 bp repeat II corresponds to nucleotides –203 to –183 (promoter central), and 21 bp repeat III corresponds to nucleotides –103 to –83 (promoter proximal) with respect to the HTLV-I transcriptional start site. (b) Sequence of the Sp1 and CREB binding site oligonucleotides utilized in the competition EMS analyses in Figure 4. Underlined nucleotides represent the CRE or Sp1 binding site. (c) Sequence of the promoter central repeat mutated oligonucleotides. Oligonucleotide designations refer to the promoter central (pc) repeat and the positions of the mutated nucleotides from the first nucleotide of the 21 bp repeat followed by the individual nucleotides native to the promoter central repeat and the mutated nucleotides inserted into the sequence of the promoter central repeat. In all cases only the sense strand is shown.

Corboy *et al*, 1992). For example, when transgenic mice were generated utilizing the LTRs from either CNS- or T cell-tropic HIV strains, expression of the reporter gene within the nervous system was detected only in mice transgenic for LTRs derived from the CNS of infected individuals with neurologic damage (Corboy *et al*, 1992). In addition, complementary studies performed to address the role of the retroviral LTR in cell type-specific viral gene expression have utilized transgenic mice containing a β-galactosidases transgene driven by an HTLV-I LTR isolated from a patient with TSP with the resultant LTR-directed expression occurring primarily within the CNS (Gonzalez-Dunia *et al*, 1992). In a parallel line of investigation, several HTLV-I proviruses isolated from ATL or TSP patients as well as HTLV-I-infected asymptomatic individuals have been isolated and sequenced (Malik *et al*, 1988; Daenke *et al*, 1990; Komurian *et al*, 1991; Sherman *et al*, 1992). At the nucleotide sequence level, although there is 95% homology between isolates, a number of the differences in nucleotide sequence between isolates lie in and around the U3 region (Daenke *et al*, 1990; Komurian *et al*, 1991). Although some evidence has been accumulated suggesting that nucleotide sequence variation within the LTR may not yet correlate in an obvious manner with disease phenotype (Gonzalez-Dunia *et al*, 1992), the precise interactions of cellular factors, the HTLV-I-encoded *trans*-activator protein, Tax, and specific LTR sequences that play a role in determining the pathogenic course after infection are not yet conclusive.

HTLV-I LTR-driven expression is dependent on *cis*-acting enhancer sequences comprised of three 21 bp imperfectly repeated elements collectively designed as the Tax-responsive element 1 (TRE-1) located within the U3 region of the LTR at positions -251 to -231 (promoter distal), -203 to -183 (promoter central), and -103 to -83 (promoter proximal) relative to the transcriptional start site (Figure 1) (Rosen *et al*, 1985; Paskalis *et al*, 1986; Shimotohno *et al*, 1986; Park *et al*, 1988). In addition, HTLV-I transcription is also dependent on a second enhancer element designated as the Tax-responsive element 2 (TRE-2) located between the promoter central and promoter proximal 21 bp repeat elements (Sodroski, 1992). Basal HTLV-I transcription is greatly enhanced by the virus-encoded regulatory protein, Tax, which mediates *trans*-activation of the HTLV-I LTR by interaction with the 21 bp repeat elements and TRE-2 (Sodroski *et al*, 1984; Felber *et al*, 1985; Seiki *et al*, 1986; Rosen *et al*, 1987; Nakamura *et al*, 1989; Sodroski, 1992). Since Tax has not been demonstrated to bind DNA directly (Niki *et al*, 1992), it has been suggested that the viral transactivator protein interacts with the HTLV-I LTR indirectly utilizing cellular intermediaries (Sodroski *et al*, 1984; Park *et al*, 1988; Giam and Xu, 1989; Fujisawa *et al*, 1991; Suzuki *et al*, 1993). While the interaction between Tax and selected transcription factors has been shown to play a critical role in both productive infection and malignant transformation, basal levels of transcription may be critical for initiation of Tax-mediated transcription. In fact, since HTLV-I does not package any detectable levels of Tax, basal transcription is solely responsible for the initial events leading to Tax-mediated transcription during productive infection and during activation of the latent proviral genome. Therefore, the characterization of the basal transcription factors interacting with the individual 21 bp repeat elements is by itself an important experimental pursuit. A number of cellular proteins which interact with the 21 bp repeat region have been identified and include ATF-CREB family members, AP-1, AP2, and Sp1 (for review see Matthews *et al*, 1992, Nyborg and Dynan, 1990; and Sodroski, 1992). For the most part, the proteins previously shown to interact with this essential viral regulatory region were isolated from cells derived from sources other than the nervous system (lymphocytes or Hela cells). Consequently, the identification and characterization of neuroglial cell factors which interact with the 21 bp repeats and possibly other regions of the LTR is essential to understand-

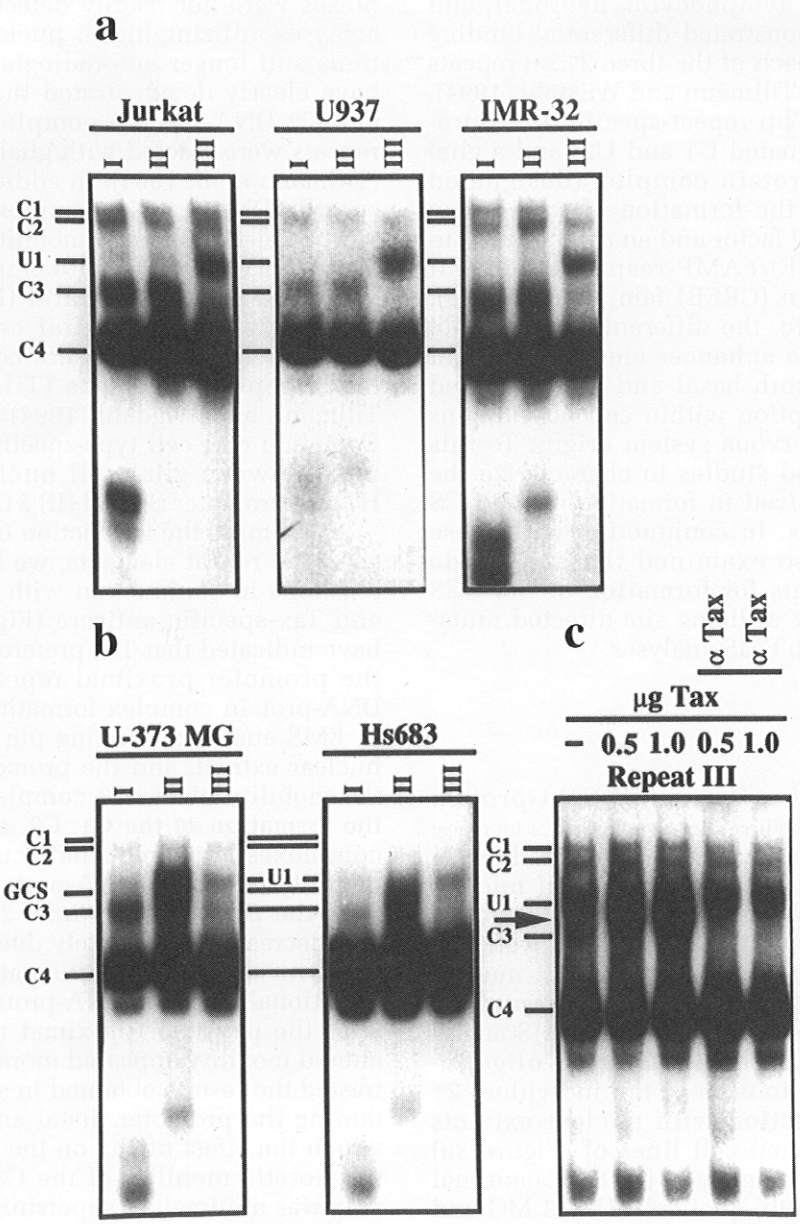

Figure 3 Identification of a glial cell-specific DNA-protein complex. Standard EMS reaction mixtures containing 9 μg of (a) lymphocytic (Jurkat), monocytic (U937), neuronal (IMR-32) or (b) glial (U-373 MG and Hs683) cell-derived nuclear extracts, 2 μg poly [d(I-C)], and 50,000 cpm of the radiolabeled 21 bp repeat I (promoter distal), II (promoter central), or III (promoter proximal) were incubated for 30 min at 30°C and subjected to electrophoresis in a 4% high-ionic-strength native polyacrylamide gel (pre-run for 1.5 h at 100 V) at 30 mA. Effect of Tax on C3 DNA-protein complex formation (c). Standard EMS reaction mixtures containing 9 μg of U-373 MG nuclear extract, 2 μg poly [d(I-C)], 50,000 cpm of the radiolabeled 21 bp repeat III (promoter proximal), and 0.5 or 1.0 μg Tax were incubated for 30 min at 30°C and subjected to electrophoresis as described. Also shown are reactions in which 9 μg of U-373 MG nuclear extract, 2 μg poly [d(I-C)], 50,000 cpm of the radiolabeled 21 bp repeat III (promoter proximal), and 0.5 or 1.0 μg Tax were incubated for 30 min at 30°C. Control or Tax antisera was added, and the reactions were incubated for an additional 30 min at 30°C and subjected to electrophoresis as described. In all cases free radiolabeled DNA probe is not shown in the actual figure. The free probe remained at an equal abundance and was present at saturating amounts at all levels of competitor DNA.

ing the molecular mechanisms involved in HTLV-I LTR-directed transcription during viral infection of glial elements within the nervous system.

While each of the 21 bp repeats have three strictly conserved domains, termed a, b, and c (Sodroski *et al*, 1984; Figure 1), these domains comprise only 13 of the 21 bp. Previous electrophoretic mobility shift (EMS) analyses utilizing double-stranded (ds) oligonucleotides homologous to each of the 21 bp repeats (Figure 2) and nuclear extracts derived from

selected cell lines of lymphocytic, neuronal, and glial origin have demonstrated differential binding of cellular factors to each of the three 21 bp repeats (Tillmann *et al*, 1994; Tillmann and Wigdahl, 1994). Specifically, both 21 bp repeat-specific DNA-protein complexes (designated U1 and U2) and a glial cell-specific DNA-protein complex (designated GCS) were detected, the formation of which may involve an Sp1-related factor and an activating transcription factor (ATF)/cAMP-response element (CRE) binding protein (CREB) family member(s), respectively. Therefore, the differential binding of cellular factors to the enhancer elements may in turn play a role in both basal and Tax-mediated LTR-directed transcription within cell populations of immune and/or nervous system origin. To this end, we have initiated studies to characterize the specific proteins involved in formation of the GCS DNA-protein complex. In conjunction with these studies, we have also examined the nucleotide sequence requirements for formation of the GCS DNA-protein complex utilizing site-directed mutagenesis in concert with EMS analyses.

Results

Formation of a glial-cell specific DNA-protein complex

EMS analyses performed with each individual 21 bp repeat and lymphocytic and glial cell nuclear extracts have resulted in the detection of 21 bp repeat-specific DNA-protein complexes as well as a DNA-protein complex specific to glial cell nuclear extracts (Tillmann *et al*, 1994; Tillmann and Wigdahl, 1994). As shown in Figure 3, EMS analyses were performed with radiolabeled oligonucleotides homologous to each of the individual 21 bp repeats in conjunction with nuclear extracts derived from established cell lines of (Figure 3a) lymphocytic (Jurkat), monocytic (U-937), neuronal (IMR-32), and (Figure 3b) glial cell (U-373 MG and Hs683) origin. When the 21 bp repeats were reacted with Jurkat, U-937, and IMR-32 nuclear extracts, four DNA-protein complexes common to each individual 21 bp repeat designated C1, C2, C3, and C4 were detected (Figure 3). In addition, a DNA-protein complex unique to the promoter proximal (III) repeat designated U1 was also detected with each nuclear extract. However, cognate DNA competition EMS analyses and EMS analyses performed in the presence of a non-specific DNA such as poly[d(I-C)] have demonstrated that the C4 DNA-protein complex is non-specific in nature (Tillmann *et al*, 1994; Tillmann and Wigdahl, 1994). When the 21 bp repeats were reacted with the nuclear extracts derived from the glial cell lines (U-373 MG and Hs683), the C3 and C4 DNA-protein complexes common to each individual 21 bp repeat were readily detected while the C1 and C2 DNA-protein complexes were not readily detectable. However, EMS analyses utilizing higher nuclear extract concentrations and longer autoradiographic exposure times have clearly demonstrated the presence of the C1 and C2 DNA-protein complexes when the 21 bp repeats were reacted with glial cell nuclear extracts (Tillmann *et al*, 1994). In addition, a glial cell type-specific DNA-protein complex designated as GCS, with an electrophoretic mobility between that of the C2 and C3 DNA-protein complexes, was detected when the promoter central (II) 21 bp repeat was reacted with either glial cell nuclear extract. Collectively, these data are consistent with previously reported results (Tillmann *et al*, 1994; Tillmann and Wigdahl, 1994) and demonstrate the formation of a cell type-specific DNA-protein complex between glial cell nuclear factors and the HTLV-I promoter central (II) 21 bp repeat.

To examine the interaction of Tax with the isolated 21 bp repeat elements, we have performed EMS reactions in conjunction with purified Tax protein and Tax-specific antisera (Figure 3c). These data have indicated that Tax preferentially interacts with the promoter proximal repeat element to affect DNA-protein complex formation. As demonstrated, in EMS analyses utilizing purified Tax, U-373 MG nuclear extract, and the promoter proximal repeat, the mobility of the C3 complex was altered while the formation of the C1, C2, and U1 DNA-protein complexes appeared unaffected. Specifically, the mobility of the C3 DNA-protein complex detected with the promoter proximal 21 bp repeat element was decreased, most likely due to the interaction of Tax with the factors participating in C3 formation. Additionally, the C3 DNA-protein complex detected with the promoter proximal repeat possessing an altered mobility appeared more abundant. This contrasted the results obtained in similar reactions containing the promoter distal and central repeats in which the effect of Tax on the abundance and electrophoretic mobility of the C3 DNA-protein complex was minimal. In supershift analyses containing the promoter proximal repeat, U-373 MG nuclear extract, purified Tax, and Tax antisera, the decrease in mobility and increase in abundance of the C3 DNA-protein complex due to purified Tax in the reactions was blocked by the addition of Tax antisera. These data suggest that Tax may alter the affinity of ATF-CREB factors for the promoter proximal repeat as previously described (Paca-Uccaralertuken *et al*, 1994).

Characterization of the protein components involved in formation of the GCS DNA-protein complex

As shown in Figure 3 and as previously reported (Tillmann *et al*, 1994), EMS analyses have demonstrated that the U1 and GCS DNA-protein complexes possess similar electrophoretic mobilities. However, competition EMS analyses utilizing unla-

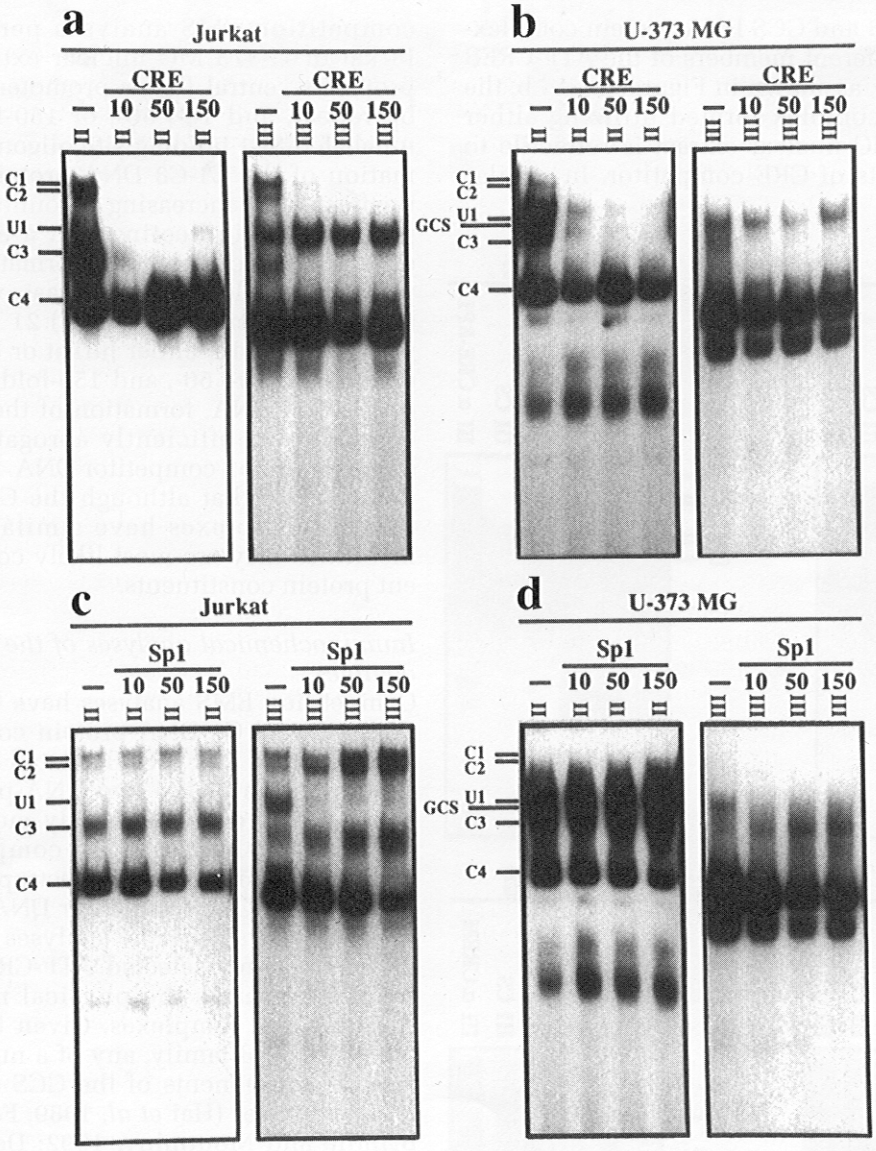

Figure 4 Competition EMS analyses with unlabeled CRE or Sp1 binding site oligonucleotides and Jurkat or U-373 MG nuclear extracts. Competition EMS reactions containing 9 μg of Jurkat (a and c) or U-373 MG (b and d) nuclear extract, 2 μg poly[d(I-C)], 50,000 cpm of the radiolabeled promoter central and promoter proximal 21 bp repeats and a 10-, 50-, or 150-fold molar excess of unlabeled CRE (a and b) or Sp1 (c and d) binding site competitor DNA were incubated for 30 min at 30°C and subjected to electrophoresis as described in Figure 3.

beled CRE and Sp1 binding site oligonucleotides have demonstrated that the two DNA-protein complexes are in fact comprised of distinct protein components (Tillmann *et al*, 1994). As demonstrated in Figure 4a, when Jurkat nuclear extract was reacted with the radiolabeled promoter central (II) or promoter proximal (III) 21 bp repeat element and 10-, 50-, or 150-fold molar excess of CRE competitor DNA, formation of the C1-C3 DNA-protein complexes was completely abrogated at less than 50-fold molar excess competitor DNA when the promoter central (II) repeat was used as the radiolabeled probe and at less than 10-fold molar excess competitor DNA when the promoter proximal (III) repeat was used as probe. This result is similar to that observed with competition EMS analyses utilizing U-373 MG nuclear extract and CRE competitor DNA (Figure 4b) and is consistent with previously reported studies (Tillmann *et al*, 1994). Although formation of the GCS DNA-protein complex detected with the promoter central (II) repeat was also inhibited by increasing amounts of CRE competitor DNA, it was necessary to utilize at least 150-fold molar excess CRE competitor DNA to obtain a comparable level of GCS DNA-protein complex abrogation. This observation suggests that the

formation of the C3 and GCS DNA-protein complexes may involve different members of the ATF-CREB family. In addition, as shown in Figure 4a and b, the U1 DNA-protein complex formed utilizing either Jurkat or U-373 MG nuclear extract is refractile to increasing amounts of CRE competitor. In similar

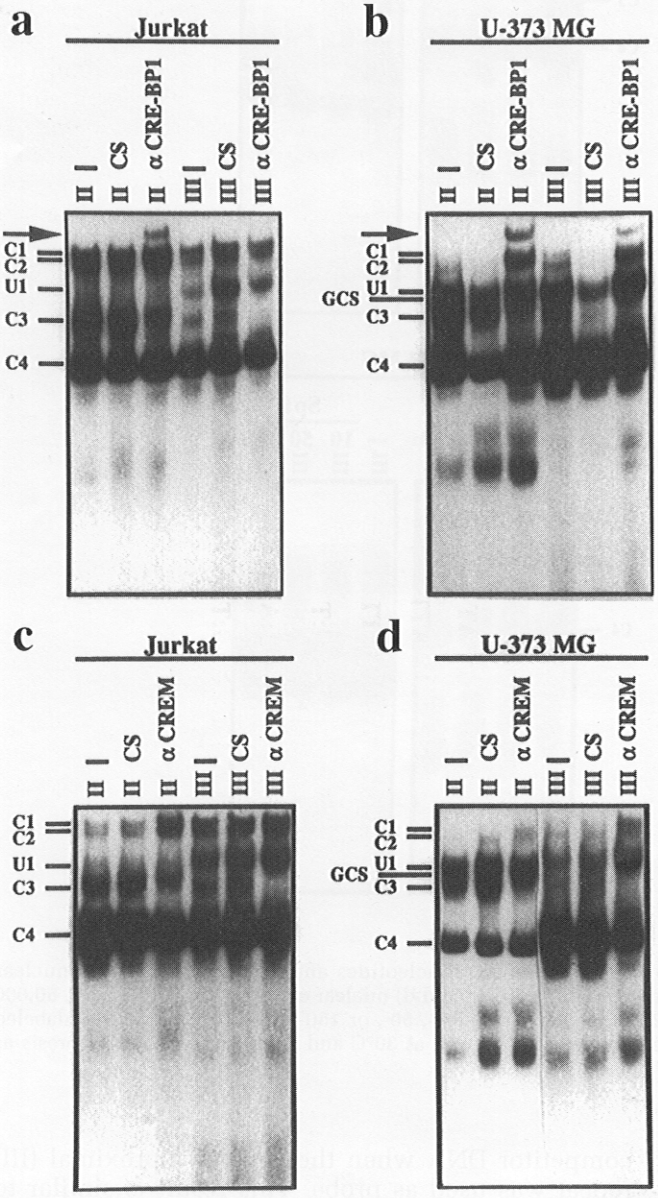

Figure 5 Supershift EMS analyses utilizing the radiolabeled promoter central or promoter distal 21 bp repeat, Jurkat or U-373 MG nuclear extract, and CREBP1 or CREM antisera. Standard EMS rections containing 9 μg Jurkat (a) or U-373 MG (b) nuclear extract, 2 μg poly[d(I-C)], and 50,000 cpm of the radiolabeled promoter central and promoter distal 21 bp repeats were incubated for 30 min at 30°C. Control sera, CREBP1, or CREM antisera was added, and the reactions were incubated for an additional 30 min at 30°C and subjected to electrophoresis as described in Figure 3. Arrows designate location of supershifted DNA-protein complexes.

competition EMS analyses performed utilizing Jurkat or U-373 MG nuclear extract, radiolabeled promoter central (II) or promoter proximal (III) 21 bp repeat, and 10-, 50-, or 150-fold molar excess unlabeled Sp1 binding site oligonucleotide, the formation of the C1-C3 DNA-protein complexes was unaffected by increasing amounts of the Sp1 competitor DNA indicating that a Sp1-related factor does not participate in the formation of these DNA-protein complexes. In contrast, when the radiolabeled promoter proximal (III) 21 bp repeat element was reacted with either Jurkat or U-373 MG nuclear extract and 10-, 50-, and 150-fold molar excess Sp1 competitor DNA, formation of the U1 DNA-protein complex was efficiently abrogated by increasing amounts of Sp1 competitor DNA. These data clearly demonstrate that although the GCS and U1 DNA-protein complexes have similar electrophoretic mobilities they are most likely comprised of different protein constituents.

Immunochemical analyses of the GCS DNA-protein complex

Competition EMS analyses have demonstrated that the GCS and C3 DNA-protein complexes are both comprised of ATF-CREB-related factors. However, the formation of the two DNA-protein complexes may involve different family members. Since the GCS and C3 DNA-protein complexes have been shown to exhibit similar electrophoretic mobilities but different CRE competitor DNA profiles, we performed supershift EMS analyses utilizing antisera directed against selected ATF-CREB factors to further characterize the physical nature of the two DNA-protein complexes. Given the complexity of the ATF-CREB family, any of a number of members may be components of the GCS and C3 DNA-protein complexes (Hai *et al*, 1989; Foulkes *et al*, 1991; Brindle and Montminy, 1992; Delmas *et al*, 1992; Laoide *et al*, 1993). Since cell type-specific ATF-CREB factors have been described (Lee, 1992), it is plausible that a glial cell-specific ATF-CREB factor(s) may be involved in GCS formation. As demonstrated utilizing the promoter central (II) or proximal (III) 21 bp repeat, Jurkat nuclear extract, and antisera directed against CREBP1, an ATF-CREB family member previously demonstrated to interact with the 21 bp repeats (Zhao and Giam, 1991), the electrophoretic mobility and relative abundance of the C1 and C2 DNA-protein complexes were unaltered (Figure 5a). In contrast, formation of the C3 DNA-protein complex was partially abrogated when utilizing the promoter central (II) repeat with a small fraction of the DNA-protein complex supershifted to a mobility less than that of the C1 DNA-protein complex (indicated by arrow) or completely abrogated when utilizing the promoter proximal (III) repeat. Utilizing either of the two repeats as radiolabeled probes in conjunction with increasing autoradiographic exposure times, it was appar-

ent that the abundance of the C1 and C2 DNA-protein complexes was increased. The observed increase is most likely due to additional supershifted DNA-protein complexes of similar mobilities or to an increased availability of target DNA as a result of abrogation of the C3 DNA-protein complex. In parallel analyses utilizing U-373 MG nuclear extract (Figure 5b), the CREBP1 antisera had no effect on GCS DNA-protein complex formation while similar results were observed for the C1-C3 DNA-protein complexes. These data indicate that CREBP1 participates in formation of the C3 DNA-protein complex with both lymphocytic and glial cell nuclear extracts but is not involved in formation of the GCS DNA-protein complex.

A CREB-related family, the CRE modulator (CREM) proteins consisting of numerous isoforms obtained by alternative splicing (Foulkes *et al*, 1991; Delmas *et al*, 1992; Laoide *et al*, 1993), has been demonstrated to interact with the HTLV-I 21 bp repeats (Suzuki *et al*, 1993). Since the CNS is a primary site of CREM expression (Foulkes *et al*, 1991; Delmas *et al*, 1992; Laoide *et al*, 1993), it was proposed that selected protein isoforms may participate in GCS DNA-protein complex formation. To examine this possibility, supershift EMS analyses were performed with a CREM antisera previously reported to specifically recognize only CREM isoforms (Delmas *et al*, 1992). As demonstrated utilizing CREM antisera and Jurkat nuclear extract (Figure 5c), the abundance of the C3 DNA-protein complex was decreased when the promoter central (II) 21 bp repeat was utilized as probe DNA or completely abrogated when the promoter proximal (III) repeat was utilized as probe DNA while the abundance of the C1 and C2 DNA-protein complexes was increased. This is most likely due to either an increased availability of target DNA as a consequence of C3 abrogation by the CREM antisera or the presence of a supershifted DNA-protein complex of a similar electrophoretic mobility to that of C1 and C2. In parallel supershift EMS analyses utilizing U-373 MG nuclear extract, the CREM antisera had no effect on the formation of the GCS DNA-protein complex while similar results with respect to the abrogation of the C3 DNA-protein complex and

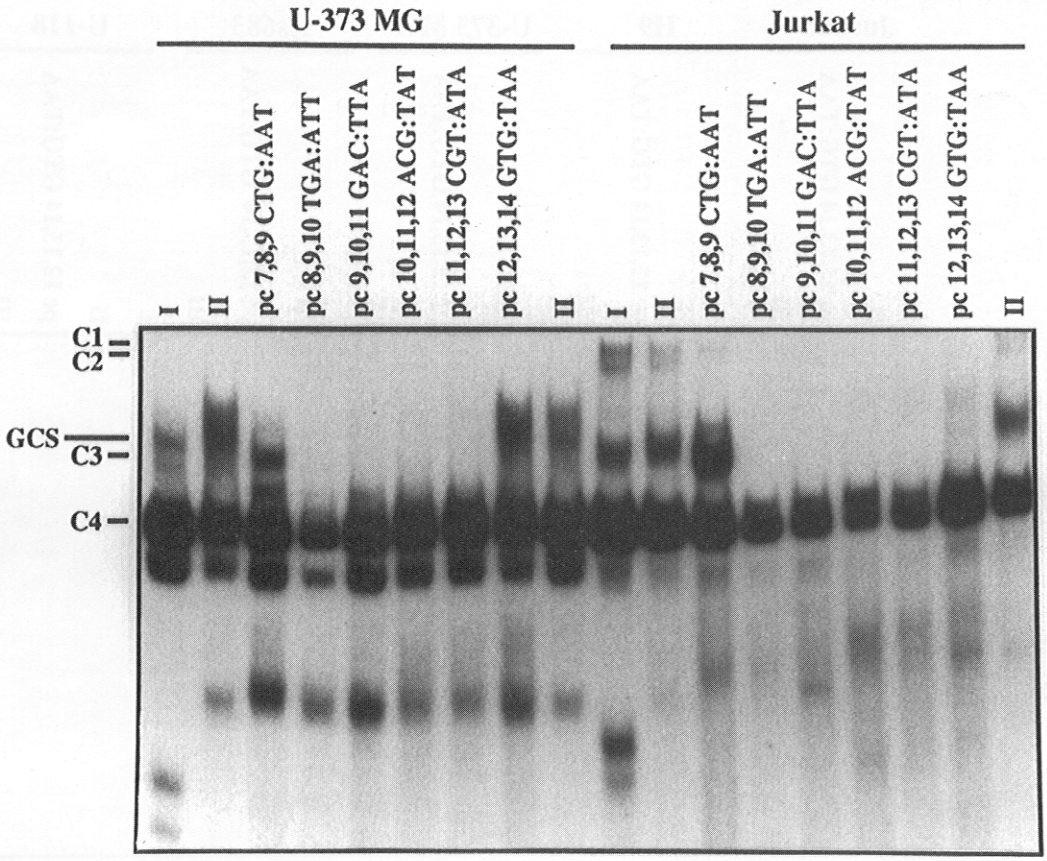

Figure 6 Effect of the mutated promoter central 21 bp repeats on DNA-protein complex formation in EMS analyses utilizing Jurkat or U-373 MG nuclear extract. Standard EMS reactions containing 9 μg of Jurkat or U-373 MG nuclear extract, 2 μg poly[d(I-C)], and 50,000 cpm of the radiolabeled mutated promoter central 21 bp repeats (Figure 2) or native 21 bp repeats were incubated for 30 min at 30°C and subjected to electrophoresis as described in Figure 3.

increase in abundance of the C1 and C2 DNA-protein complexes were observed. Collectively, these data demonstrate a novel interaction between the ATF-CREB family members CREM and CREBP1 with the HTLV-I 21 bp repeats in the formation of the C3 DNA-protein complex. Although the C1 and C2 DNA-protein complexes were not supershifted by CREBP1 or CREM antisera, similar supershift EMS analyses utilizing specific antisera demonstrated that these DNA-protein complexes are comprised of the transcription factors ATF-1 and ATF-2 (Wessner *et al*, 1995, manuscript in preparation). Although the GCS DNA-protein complex is comprised of factors that recognize the CRE binding site consensus sequence, supershift analyses have indicated that the ATF-CREB factors CREBP1 and CREM (Figure 5), CREB2 (data not shown) as well as ATF-1 and ATF-2 (Wessner *et al*, 1995, manuscript in preparation) are not involved in the formation of this DNA-protein complex. In addition, supershift EMS analyses utilizing antisera directed against AP-1 or its respective components (Fos/Jun), factors which recognize a core consensus sequence highly homologous to the CRE binding site, demonstrated that these factors are also not involved in the formation of the GCS DNA-protein complex (Wessner *et al*, 1995, manuscript in preparation).

Nucleotide sequence requirements for formation of the GCS DNA-protein complex

As determined by oligonucleotide competition EMS analyses, the factor(s) participating in GCS DNA-protein complex formation have been demonstrated to have a lower affinity for the CRE consensus sequence than those participating in C1–C3 formation. Consequently, to determine the precise nucleotide sequence requirements for formation of the GCS and C1–C3 DNA-protein complexes, site-directed mutagenesis in which overlapping sets of three nucleotides replacing selected nucleotides of the CRE binding site in the promoter central repeat (Figure 2b) were performed. When the mutated promoter central (pc) repeat pc7,8,9 CTG-AAT, in which the first three nucleotides (CTG) of the CRE binding site (at nucleotides numbered 7, 8, and 9 beginning from the first nucleotide of the promoter

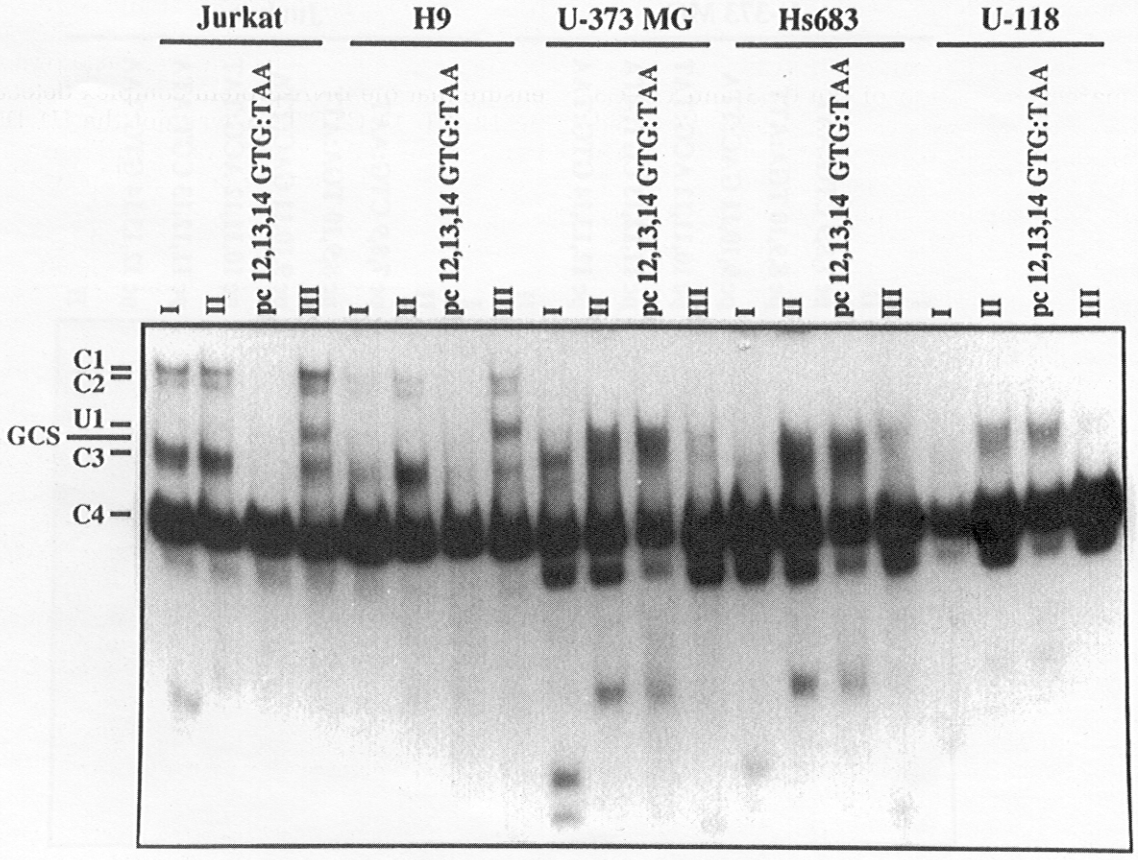

Figure 7 EMS analyses utilizing pc14,15,16 GTG:TAA and nuclear extracts derived from selected cell lines of T cell of glial cell origin. Standard EMS reactions containing 9 μg of Jurkat, H9, U-373 MG, Hs683, or U-118 nuclear extract, 2 μg poly [d(I-C)], and 50,000 cpm of pc14,15,16 GTG:TAA or native 21 bp repeats were incubated for 30 min at 30°C and subjected to electrophresis as described in Figure 3.

central repeat) were mutated to ATT, was reacted with U-373 MG or Jurkat nuclear extract, formation of the C1 and C2 and GCS DNA-protein complexes was completely abrogated (Figure 6). Initial EMS analyses utilizing pc7,8,9 CTG:AAT resulted in the formation of a DNA-protein complex with an electrophoretic mobility similar to that of the C3 DNA-protein complex. However, competition EMS analyses utilizing CRE binding site competitor DNA had no effect on the formation of this DNA-protein complex suggesting that the DNA-protein complex detected with pc7,8,9 CTG:AAT was not related to the C3 DNA-protein complex (data not shown). Additionally, formation of a DNA-protein complex with a mobility greater than the C3 DNA-protein complex but less than the C4 DNA-protein complex was detected when pc7,8,9 CTG:AAT was reacted with U-373 MG nuclear extract. The nature of these novel DNA-protein complexes detected with pc7,8,9 CTG:AAT is currently under investigation. However, inspection of the pc7,8,9 CTG:AAT sequence revealed the presence of a binding site for CAAT-related factors indicating that this DNA-protein complex may be due to an interaction of the mutated target sequence with a CAAT-related protein. When pc8,9,10 TGA:AAT, pc9,10,11 GAC:TTA, pc10,11,12 ACG:TAT, and pc11,12,13 CGT:ATA were reacted with U-373 MG or Jurkat nuclear extract, formation of the GCS and C1–C3 DNA-protein complexes was completely abrogated (Figure 6). In contrast, when pc12,13,14 GTG:TAA was reacted with U-373 MG nuclear extract, formation of a DNA-protein complex with similar electrophoretic mobility to the GCS complex was detected while formation of the C1–C3 DNA-protein complexes was completely abrogated. The DNA-protein complex detected with pc12,13,14 GTG:TAA and U-373 MG nuclear extract was not the C3 DNA-protein complex since its formation was not detected in reactions utilizing Jurkat nuclear extract. As expected, when each of the mutated central (II) 21 bp repeats were utilized in EMS analyses with Jurkat or U-373 MG nuclear extract, formation of the non-specific C4 DNA-protein complex was unaffected, a result similar to that previously reported with site-directed mutagenesis of the promoter proximal (III) repeat (Wessner *et al*, 1995, submitted for publication). To ensure that the results obtained with oligonucleotide pc12,13,14 GTG:TAA and U-373 MG nuclear extract was representative of other glial cell lines and not specific to the U-373 MG cell line, EMS analyses utilizing nuclear extracts derived from selected T cell (Jurkat and H9) and glial cell (U-373 MG, Hs683, and U-118) lines and pc12,13,14 GTG:TAA were performed (Figure 7). As shown, results similar to those obtained utilizing U-373 MG nuclear extract were observed when pc 12,13,14 GTG:TAA was reacted with nuclear extracts derived from two independent glial cell lines (Hs683 and U-118). Furthermore, results similar to those obtained utilizing Jurkat nuclear extract were observed when pc 12,13,14 GTG:TAA was reacted with nuclear extract from an additional lymphocyte cell line (H9). These data further substantiate the glial cell-specific nature of the DNA-protein complex detected with pc 12,13,14 GTG:TAA.

Characterization of the components participating in formation of the DNA-protein complex detected with pc 12,13,14 GTG:TAA and glial cell nuclear extract

To further demonstrate that the DNA-protein complex detected with pc 12,13,14 GTG:TAA is related to the GCS DNA-protein complex, competition EMS analyses were performed utilizing radiolabeled pc 12,13,14 GTG:TAA, U-373 MG or Hs683 nuclear extract, and 10-, 50-, and 150-fold molar excess of unlabeled CRE binding site DNA. Since the Sp1-associated U1 DNA-protein complex detected with the promoter proximal (III) repeat has a similar electrophoretic mobility to the GCS DNA-protein complex and the DNA-protein complex detected with pc 12, 13,14 GTG:TAA, competition EMS analyses utilizing radiolabeled pc 12,13,14 GTG:TAA, U-373 MG or Hs683 nuclear extract, and a 10-, 50-, and 150-fold molar excess of unlabeled CRE and Sp1 binding site oligonucleotide were also performed to ensure that the DNA-protein complex detected with pc 12, 13, 14 GTG:TAA was not the U1 DNA-protein complex. As demonstrated in Figure 8, formation of the DNA-protein complex detected with pc

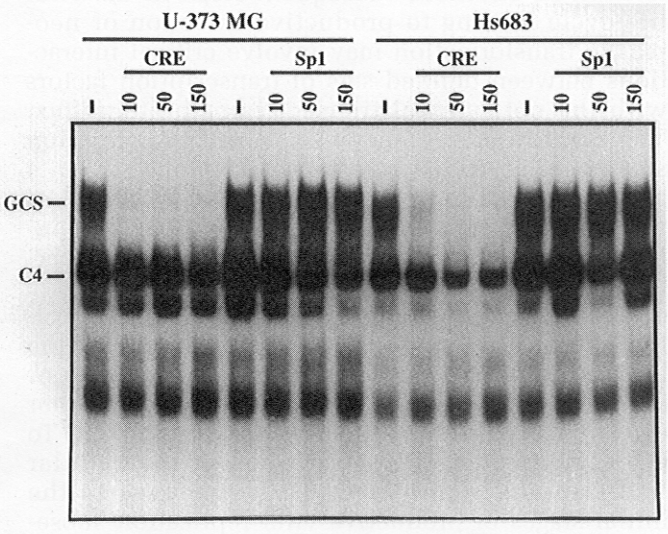

Figure 8 Competition EMS analyses with utilizing radiolabeled pc14,15,16 GTG:TAA, U-373 MG or Hs683 nuclear extract, and unlabeled CRE or Sp1 binding site oligonucleotides. Competition EMS reactions containing 9 µg of U-373 MG or Hs683 nuclear extract, 2 µg poly[d(I-C)], 50,000 cpm of pc 14,15,16 GTG:TAA, and a 10-, 50-, or 150-fold molar excess of unlabeled CRE or Sp1 binding site oligonucleotide for 30 min at 30° and subjected to electrophoresis as described in Figure 3.

12,13,14 GTG:TAA and U-373 MG or Hs683 nuclear extract was abrogated by increasing amounts of CRE binding site DNA. In contrast, increasing amounts of Sp1 binding site DNA had no effect on formation of the DNA-protein complex detected with pc 12,13,14 GTG:TAA and U-373 MG or Hs683 nuclear extract. These results further substantiate the glial cell-specific nature of the DNA-protein complex detected with the mutated promoter central (II) 21 bp repeat element pc 12,13,14 GTG:TAA since ATF-CREB-related factors and not Sp1-related factors participate in the formation of this DNA-protein complex.

Discussion

Regulation of transcription in a wide variety of eukaryotic cell populations during development, organ- and cell type-specific gene expression, and oncogenic transformation and tumorigenesis involves intricate interactions between a large number of transcription factor families. The inherent complexity of these interactions is magnified during infection of a host cell with any of a number of human viruses. The retroviral life cycle may involve one or more pathways that include productive viral replication leading to host cell death, viral latency or persistence, and neoplastic transformation. With respect to HTLV-I infection, viral transcription is critically dependent on the host cell transcriptional machinery for initiation of viral gene expression subsequent to entry into the target cell and during the early stages of virus activation from latency. In addition, subsequent steps in the viral life cycle leading to productive infection or neoplastic transformation may involve critical interactions between defined sets of transcription factors with the potent viral transactivator protein, Tax. The absence or presence of selected transcription factors, either individually or in combination, in a given cell type may impact on the overall level of viral genomic expression in that specific cell population. Furthermore, the transcription factor repertoire within given cell populations may also have direct bearing on Tax-mediated upregulation of cellular genes that may impact on conversion of the host cell from a normal to malignant phenotype or on cellular processes that may be involved either directly or indirectly in neurologic dysfunction. To begin systematic characterization of the cellular transcription factors that may be involved in the initial stages of productive viral replication subsequent to virus entry or activation of an endogenous infection in either the immune or nervous system, we have performed EMS analyses in which the HTLV-I individual 21 bp repeats were reacted with nuclear extracts derived from cells of lymphocytic and neuroglial cell origin. These studies have now provided the foundation for future studies to identify and characterize the immune or neuroglial factors that may interact with the HTLV-I Tax protein to boost viral gene expression or perturbate cellular function. These studies have resulted in the detection of several groups of DNA-protein complexes comprised of a number of different members of several transcription factor families. The first group consists of a series of DNA-protein complexes, designated C1–C3, which are comprised of members of the ATF-CREB transcription factor family and are common to each cell type examined and can be detected with each of the three isolated 21 bp repeat elements. The second group consists of two DNA-protein complexes, designated U1 and U2, which are also common to each cell type examined but are comprised of members of the Sp1 family of transcription factors and can only be detected with the promoter proximal (III) 21 bp repeat. The third group of DNA-protein complexes is represented by the GCS DNA-protein complex which specifically forms with glial cell nuclear extract and the promoter central (II) 21 bp repeat and is comprised of ATF-CREB family members which differ from those comprising the C1–C3 DNA-protein complexes.

Since the C1–C3 DNA-protein complexes detected with the individual repeats and both lymphocyte and glial cell nuclear extracts migrate with identical electrophoretic mobilities, similar proteins capable of interacting with the CRE binding site of the 21 bp repeats to form the C1–C3 DNA-protein complexes appeared to be present in the two cell types. In this regard, we present evidence that the ATF-CREB-related factors participating in the formation of the C1–C3 DNA-protein complexes detected with glial cell nuclear extract are indeed identical or highly related to those detected with lymphocyte nuclear extract and further substantiate the glial cell type-specific nature of the ATF-CREB-related factors participating in formation of the GCS DNA-protein complex. Based on supershift EMS analyses, the C3 DNA-protein complex contains the CREBP1 protein which has previously been shown to interact with the 21 bp repeats (Zhao and Giam, 1991). In addition, an isoform of the newly characterized ATF-CREB subgroup, the cyclic AMP response element modulator (CREM), a protein, also interacts with each of the individual 21 bp repeats to participate in formation of the C3 DNA-protein complex.

Although the C1–C3 DNA-protein complexes are common to each cell type examined, the relative abundance of these DNA-protein complexes was strikingly different between selected cell populations. Therefore, the transcription factor repertoire within given cell populations may lead to cell type-specific interactions with the isolated 21 bp repeats which may have significant impact on control of viral transcription. Specifically, with each glial cell line examined, the abundance of the C1–C2 DNA-protein complexes was much lower than the abundance of these complexes detected with lymphocytic, monocytic, or neuronal nuclear extracts while

the abundance of the C3 DNA-protein complex remained relatively constant with respect to all nuclear extracts examined. Supershift EMS analyses have demonstrated that the C1 DNA-protein complex contains ATF-1 and ATF-2 while the C2 DNA-protein complex contains ATF-2 and Jun (Wessner *et al*, 1995, manuscript in preparation). Since Tax has been demonstrated to enhance the binding activity of selected ATF-CREB family members including ATF-1 and ATF-2 (Armstrong *et al*, 1993), Tax *trans*-activation of HTLV-I transcription may be mediated, in part, through the increased binding of these factors to the 21 bp repeat elements. Therefore, it is possible that the reduced abundance of the C1–C3 DNA-protein complex and associated transcription factors observed with glial cell nuclear extracts may be linked in some manner to less productive HTLV-I replication in neuroglial cell populations.

In addition to variations in abundance of DNA-protein complex formation between the cell types examined, variations in abundance of the C3 DNA-protein complex common to all cell type examined occurs between the individual 21 bp repeats. We demonstrate that, although the C3 DNA-protein complex forms with each individual 21 bp repeat, the abundance of the C3 DNA-protein complex detected with the promoter proximal (III) repeat is greatly reduced as compared to the abundance of the C3 DNA-protein complexes detected with the promoter distal (I) and central (II) repeats. It is interesting to note that two 21 bp repeat-specific DNA-protein complexes, designated U1 and U2, also form with the promoter proximal (III) repeat and contain the transcription factor Sp1 (Tillmann *et al*, 1994; Wessner *et al*, 1995, submitted for publication). Similar site-directed mutagenesis performed with the promoter proximal (III) repeat demonstrated that a competitive interaction between the factors comprising the U1 and U2 DNA-protein complex and the factors comprising the C1–C3 DNA-protein complexes since mutations that increase the abundance of one type of DNA-protein complex reduce the abundance of the other and *vice versa* (Wessner *et al*, 1995, submitted for publication). Therefore, the complex interaction between ATF-CREB-related factors and Sp1 with the promoter proximal (III) repeat may lead to a differential functional activity of the promoter proximal (III) repeat as compared to the promoter distal (I) and central (II) repeats.

As outlined above, we have demonstrated that the C3 DNA-protein complex is comprised of at least the ATF-CREB family members, CREBP1 and CREM. The CREM gene can encode both activators and repressors of cAMP-induced transcription based on patterns of alternative splicing. To date, approximately seven CREM isoforms have been identified (Delmas *et al*, 1992). Each isoform possesses different structural domains generated by exon shuffling to produce protein products capable of interaction with CRE binding sites and heterodimerization with other ATF-CREB family members. Although the antisera utilized in these studies specifically recognizes CREM proteins from other ATF-CREB factors, it cannot discern between CREM isoforms (Delmas *et al*, 1992). It is possible that a CREM activator or repressor (or activator-repressor combinations) may interact with CREBP1 to participate in the formation of the C3 DNA-protein complex and subsequently modulate transcription. Recent evidence demonstrates that both CREM activators and repressors homodimerize and readily heterodimerize with other CREM isoforms and ATF-CREB factors (Laoide *et al*, 1993). In addition, it appears that the CREM repressor proteins exert their negative effects on transcription through the formation of antagonistic dimers with ATF-CREB factors and CREM activator proteins (Laoide *et al*, 1993). Since the CREM antisera utilized in these studies cannot discern between CREM activators or repressors, it is not possible to discern which CREM isoform may interact with CREBP1 to participate in the formation of the C3 DNA-protein complex. It is interesting to speculate about the role of CREM in HTLV-I transcription mediated through the 21 bp repeats. Recent evidence suggests that Tax can interact with a CREM isoform and participate in DNA-protein complex formation in EMS analyses utilizing five tandem copies of the promoter central (II) repeat (Suzuki *et al*, 1993). Therefore, if the CREM isoform participating in formation of the C3-DNA-protein complex detected with each individual 21 bp repeat is an activator of transcription, Tax may interact with the CREM protein to facilitate viral transcription from the 21 bp enhancer region. Conversely, the CREM isoform participating in formation of the C3 DNA-protein complex may be a repressor of transcription and through an interaction with Tax act to sequester the viral transactivator protein and limit its *trans*-activating potential. In support of this hypothesis, we have generated evidence that suggests that the promoter proximal (III) 21 bp repeat is more transcriptionally active with respect to Tax-mediated *trans*-activation of the LTR and that purified Tax protein preferentially interacts with the promoter proximal (III) repeat (Figure 3, Wessner and Wigdahl unpublished data). Since the C3 DNA-protein complex is always detected at a lower abundance than those detected with the promoter distal (I) and central (II) repeats, one could speculate that the increased activity of the promoter proximal (III) repeat may be due to a lower abundance of a CREM repressor protein interacting with the promoter proximal (III) repeat.

Although most, if not all, of the ATF-CREB related factors participating in formation of the C1–C3 DNA-protein complexes have been ascertained, the identity of the ATF-CREB-related factors participating in the formation of the GCS DNA-protein com-

plex have yet to be identified. As we have shown, the major glial cell-specific difference in DNA-protein complex formation with the individual 21 bp repeats is the formation of the GCS DNA-protein complex detected with the promoter central (II) 21 bp repeat. Since the CRE consensus sequence is highly homologous to the consensus sequence for AP-1 and its associated factors, Fos and Jun, and since ATF-CREB and AP-1 factors occasionally interact with overlapping binding sites within promoter elements, the GCS DNA-protein complex may be comprised of AP-1-related factors. To this end, supershift EMS analyses were performed utilizing AP-1 specific antisera (data not shown). As with the supershift analyses utilizing the selected ATF-CREB antisera, formation of the GCS DNA-protein complex was unaffected by AP-1, Fos, or Jun antisera. Since a wide variety of ATF-CREB-related factors (CREBP1, CREB2, CREM, ATF-1, and ATF-2) and AP-1 factors (AP-1, Fos and Jun) can be ruled out as participants in GCS DNA-protein complex formation, it appears that a novel transcription factor may be present in glial cells that is capable of interaction with the promoter central (II) 21 bp repeat. Furthermore, we performed mixing experiments with glial cell and lymphocyte nuclear extracts in which one extract was kept at a constant protein concentration while the other was added in increasing amounts well over the constant concentration of the other to address the statement concerning the possibility of the presence of a repressor in the lymphocyte nuclear extract and an activator in the glial cell nuclear extract (data not shown). These experiments demonstrated that the GCS DNA-protein complex readily formed in the presence of large amounts of lymphocyte nuclear extract (data not shown). These data suggest that formation of the GCS DNA-protein complex is not due to the absence of an activating protein necessary for recruitment of the factors into GCS DNA-protein complex formation in lymphocyte nuclear extract which is present in glial cell nuclear extract or to the presence of a repressor protein inhibiting GCS DNA-protein complex formation in lymphocyte nuclear extract which is absent in glial cell nuclear extract (data not shown). Site-directed mutagenesis of the promoter proximal (III) repeat has demonstrated that the DNA-binding specificity of the ATF-CREB related factors participating in formation of the GCS DNA-protein complex for the CRE binding site in the promoter central (II) repeat are distinct from those participating in formation of the C1–C3 DNA-protein complexes further indicating the unique nature of this ATF-CREB-related factor. Specifically, the mutated promoter central (II) repeat pc12,13,14 GTG:TAA, when reacted with glial cell nuclear extract, resulted in the selective formation of the GCS DNA-complex and the specific abrogation of the C1–C3 DNA-protein complexes. Utilization of a multimerized form of this oligonucleotide to probe a glial cell-derived cDNA expression library will facilitate the characterization and cloning of the glial cell-specific factors participating in GCS formation while eliminating potential difficulties which may arise when using a radiolabeled native promoter central (II) repeat which forms both the GCS and C1–C3 DNA-protein complexes. Based on the absence of immunoreactivity of a number of antisera directed against known transcription factors with the GCS DNA-protein complex, it is likely that complex is derived of a potentially novel transcription factor(s) specifically expressed in neuroglial cell populations. The subsequent cloning and characterization of this factor will permit the examination of the functional contribution of this factor to Tax-mediated HTLV-I expression and gain insight to the differential viral regulation observed between cell types of immune and nervous system origin.

Materials and methods

Cell culture

Human T cell lymphocytic cell lines utilized in these studies include Jurkat (ATCC No. TIB 152) and H9 (National Institutes of Health AIDS Research and Reference Reagent Program). Both T cell lines and the U-937 monocytic cell line (ATCC CRL 1593) were cultured and maintained at 37°C in 5% CO_2 in RPMI 1640 medium (Gibco BRL, MD). Human glial cell lines utilized in these studies include U-373 MG (ATCC No. HTB 17), Hs683 (ATCC No. HTB 138), and U-118 (ATCC No. HTB 15). All glial cell lines and the human neuronal cell line IMR-32 (ATCC CCL 127) were cultured and maintained at 37°C in Eagle minimal essential medium (Gibco BRL, MD). All media formulations were supplemented with 10% heat-inactivated fetal calf serum, 100 U ml^{-1} penicillin, 100 µg ml^{-1} streptomycin, 2mM L-glutamine, and 0.075% $NaHCO_3$.

Nuclear extracts

Nuclear extracts of all cell lines were prepared as described in Dignam *et al*, 1983 with minor modifications. Briefly, cells were collected from cultures during logarithmic growth and nuclei were isolated utilizing hypotonic buffer (10 mM HEPES [pH 7.9, 4°C], 1.5 mM $MgCl_2$, 10 mM KCl, 0.2 mM PMSF, 0.5 mM DTT) and a Dounce homogenizer (type B) (Bellco, NJ) at 4°C. Nuclear proteins were extracted at 4°C utilizing high-salt buffer (20 mM HEPES [pH 7.9, 4°C], 25% glycerol, 1.5 mM $MgCl_2$, 700 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT) 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 PMSF, 0.5 mM DTT) for 5 h at 4°C. Precipitated proteins and cellular debris were removed by centrifugation at 21,000 × *g* for 30 min at 4°C. Protein quantitation of the

nuclear extracts was performed utilizing the Bio-Rad protein assay which is based on the Bradford dye-binding procedure (Bradford, 1976). The nuclear extracts were frozen in liquid nitrogen and stored at –80°C at a concentration of 3 mg ml^{-1}.

Oligonucleotide synthesis and radiolabeling
Complementary single-stranded oligonucleotides with native HTLV-I LTR flanking sequences corresponding to each of the three individual 21 bp repeat elements and the mutated promoter central (II) repeat elements were synthesized with 5′-extensions. Single-stranded complementary oligonucleotides were annealed by heating for 10 min at 90°C and cooling to room temperature. CREB and Sp1 binding site oligonucleotides utilized in competition EMS analyses were obtained from Stratagene. The sequences of all oligonucleotides utilized in these studies are illustrated in Figure 2. The double-stranded (ds) oligonucleotides with 5′-extensions were labeled utilizing [α-^{32}P]-dCTP and the Klenow fragment of DNA polymerase I.

EMS analyses
Standard EMS analyses were performed essentially as described by Garner and Revzin (1981). Approximately 50,000 cpm of radiolabeled ds-oligonucleotide (0.1–1.0 ng) was reacted with 6 to 9 µg of protein and 2 µg of poly[d(I-C)] in a total reaction volume of 15 µl for 30 min at 30°C. Following incubation, 2 µl of loading buffer (50% glycerol, 0.1 M EDTA [pH 8.0], 0.1% bromophenol blue, and 0.1% xylene cyanol) was added and each reaction was subjected to electrophoresis in a 4% high-ionic-strength native polyacrylamide gel (pre-run for 1.5 h at 100 V) at 30 mA. The polyacrylamide gels were dried under vacuum at 80°C for 1.5 h prior to autoradiography. In EMS analyses containing purified Tax, 0.5 or 1.0 µg of Tax (kindly provided by Dr Chou-Zen Giam, Case Western Reserve University) was added simultaneously with the other components of the standard EMS reactions and subjected to electrophoresis. For the EMS reactions containing Tax antisera (kindly provided by Dr Chou-Zen Giam, Case Western Reserve University), standard EMS reactions containing 0.5 or 1.0 µg of Tax were performed and incubated for 30 min at 30°C. After allowing the DNA-protein complexes to form for 30 min at 30°C, 1 µl of either Tax or an equivalent amount of irrelevant antisera was added to the reactions. The reactions were incubated for an additional 30 min at 30°C and subjected to electrophoresis. In competition EMS analyses, unlabeled competitor oligonucleotides were simultaneously incubated with nuclear extract and radiolabeled probe DNA for 30 min at 30°C and subjected to electrophoresis. For supershift EMS analyses (Paca-Uccaralertuken *et al*, 1994), the radiolabeled oligonucleotide was reacted with nuclear extract as described above. After allowing the DNA-protein complexes to form for 30 min at 30°C, 1 µl of either CREBP1 (kindly provided by Peter Beimling, Max-Planck-Institut für Molekulare Genetik, Berlin, Germany) or CREM (kindly provided by Paolo Sassone-Corsi, Institut de Chimie Biologique, Strasbourg, France) antisera or an equivalent amount of irrelevant antisera was added to the reactions. The reactions were incubated for an additional 30 min at 30°C and subjected to electrophoresis.

Acknowledgements

The studies performed in the laboratory of BW were supported by a grant from the National Cancer Institute (Public Health Service grant CA 54559). We acknowledge Joseph J Kenny and Fred Krebs for valuable discussions and critical reading of the manuscript. All oligonucleotides utilized in these studies were synthesized in the Macromolecular Core Facility at The Pennsylvania State University College of Medicine.

References

Akagi T, Hoshida Y, Yoshino T, Teramoto N, Kondo E, Hayashi K, Takahashi K (1992). Infectivity of human T-cell leukemia virus type I to human nervous tissue cells *in vitro*. *Acta Neuropathol* **84**: 147–52.

Armstrong AP, Franklin AA, Uittenbogaard MN, Giebler HA, Nyborg J (1993). Pleiotropic effect of the human T-cell leukemia virus Tax protein on the DNA-binding activity of eukaryotic transcription factors. *Proc Natl Acad Sci USA* **90**: 7303–7307.

Brindle PK, Montminy MR (1992). The CREB family of transcription activators. *Curr Opin Genet Dev* **2**: 199–204.

Bhagavati S, Ehrlich G, Kula RW, Kwok S, Sninsky J, Udani V, Poiesz BH (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.

Bhigjee AI, Wiley CA, Wachsman W, Amenomori T, Pirie D, Bill PLA, Windsor I (1991). HTLV-I-associated myelopathy: Clinicopathologic correlation with localization of provirus to the spinal cord. *Neurology* **41**: 1990–1992.

Bradford MM (1976). A rapid sensitive method for the quantitation of microgram quantities of proteins utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254.

Celander D, Haseltine WA (1984). Tissue-specific transcription preference as a determinant of cell tropism and leukaemogenic potential of murine retroviruses. *Nature* (London) **312**: 159–162.

Chatis PA, Holland CA, Hartley JW, Rowe WP, Hopkins N (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.

Corboy JR, Buzy JM, Zink MC, Clements JE (1992). Expression directed from HIV long terminal repeats in the central nervous system of transgenic mice. *Science* **258**: 1804–1808.

Daenke S, Nightendale S, Cruickshank JK, Bangham CRM (1990). Sequence variants of human T-cell lymphotropic virus type I from patients with tropical spastic paraparesis and adult T-cell leukemia do not distinguish neurological from leukemic isolates. *J Virol* **64**: 1278–1282.

Delmas V, Laoide BM, Masquilier D, de Groot RP, Foulkes NS, Sassone-Corsi P (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.

Dignam JD, Lebovitz RM, Roeder RG (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11**: 1475–1489.

Felber BK, Paskalis H, Kleinman-Ewing C, Wong-Staal F, Pavlakis GN (1985). The pX protein of HTLV-I is a transcriptional activator of its long terminal repeats. *Science* **229**: 675–679.

Foulkes NS, Borelli E, Sassone-Corsi P (1991). CREM gene: Use of alternative DNA-binding domains generates multiple antagonists of cAMP-induced transcription. *Cell* **64**: 739–749.

Fujisawa J-I, Toita M, Yoshimura T, Yoshida M (1991). The indirect association of human T-cell leukemia virus *tax* protein with DNA results in transcriptional activation. *J Virol* **65**: 4525–4528.

Garner MM, Revzin A (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.

Gessain A, Vernant JC, Maurs L, Barin F, Gout O, Celander A, de The G (1985). Antibodies to human T-lymphotropic virus type I in patients with tropical spastic paraparesis. *Lancet* **24**: 407–410.

Giam C-Z, Xu Y-L (1989). HTLV-I *tax* gene product activates transcription via pre-existing cellular factors and cAMP responsive element. *J Biol Chem* **264**: 15236–15241.

Gonzalez-Dunia D, Grimber DG, Briand P, Brahic M, Ozden S (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.

Hai T-Y, Liu F, Couko W, Green M (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.

Iannone R, Sherman MP, Rodgers-Johnson PEB, Beilke MA, Mora CA, Amin RM, Tinsley SR, Papsidero LD, Poiesz BJ, Gibbs CJ (1992). HTLV-I DNA sequences in CNS tissue of a patient with tropical spastic paraparesis and HTLV-I myelopathy. *J Acquir Immune Defic Syndr* **5**: 810–816.

Kira J-I, Itoyama Y, Koyanagi Y, Tateishi J, Kiskikawa M, Akizuki S-I, Kobayashi I, Toki N, Sueishi K, Sato H, Sakaki Y, Yamamoto N, Goto I (1992). Presence of HTLV-I proviral DNA in central nervous system of patients with HTLV-I-associated myelopathy. *Ann Neurol* **31**: 39–45.

Komurian F, Pelloquin F, de The G (1991). *In vivo* genomic variability of human T-cell leukemia virus type I depends more upon geography than upon pathologies. *J Virol* **65**: 3770–3778.

Laoide BM, Foulkes NS, Schlotter F, Sassone-Corsi P (1993). The functional versatility of CREM is determined by its modular structure. *EMBO J* **12**: 1179–1191.

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.

Li Y, Golemis E, Hartley JW, Hopkins N (1987). Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides. *J Virol* **61**: 693–700.

Malik K, Even J, Karpas A (1988). Molecular cloning and complete nucleotide sequence of an adult T-cell leukeamia virus/human T-cell leukaemia virus type I (ATLV/HTLV-I) isolate of Caribbean origin: relationship to other members of the ATLV/HTLV-I subgroup. *J Gen Virol* **69**: 1695–1710.

Matthews M-AH, Markowitz R-B, Dynan WS (1992). *In vitro* activation of transcription by the human T-cell leukemia virus type I *tax* protein. *Mol Cell Biol* **12**: 1986–1996.

Nakamura M, Niki M, Ohtani K, Sugamura K (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.

Niki M, Ohtani K, Nakamura M, Sugamura K (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.

Nyborg JK, Dynan WS (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.

Paca-Uccaralertuken S, Zhao L-J, Adya N, Cross JV, Cullen BR, Boros IM, Giam C-G (1994). *In vitro* selection of DNA elements highly responsive to the human T-cell lymphotropic virus type I transcriptional activator, Tax. *Mol Cell Biol* **14**: 456–462.

Paquette Y, Hanna Z, Savard P, Brousseau R, Robitaille Y, Jolicoeur P (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.

Park RE, Haseltine WA, Rosen CA (1988). A nuclear factor is required for transactivation of HTLV-I gene expression. *Oncogene* **3**: 275– 279.

Paskalis H, Felber BK, Pavlakis GN (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.

Popovic M, Sarin PS, Robert-Gurroff M, Kalyanaraman VS, Mann D, Minowada J, Gallo RC (1983). Isolation and transmission of human retrovirus (Human T-cell leukemia virus). *Science* **219**: 856–859.

Rosen CA, Haseltine WA, Lenz J, Ruprecht R, Cloyd MW (1985). Tissue selectivity of murine leukemia virus infection is determined by long terminal repeat

sequences. *J Virol* **55**: 862–866.

Rosen CA, Park R, Sodroski JG, Haseltine WA (1987). Multiple sequence elements are required for regulation of human T-cell leukemia virus gene expression. *Proc Natl Acad Sci USA* **84**: 4919–4923.

Rosen CA, Sodroski JG, Haseltine WA (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.

Seiki M, Inoue J-I, Takeda T, Yoshida M (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.

Sherman MP, Saksena NK, Dube DK, Yanagihara R, Poiesz BJ (1992). Evolutionary insights on the origin of human T-cell lymphoma/leukemia virus type I (HTLV-I) derived from sequence analysis of a new HTLV-I variant from Papua New Guinea. *J Virol* **66**: 2556–2563.

Shimotohno K, Takano M, Teruuchi T, Miwa M (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.

Sodroski JG, Rosen CA, Haseltine WA (1984). Trans-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. *Science* **225**: 381–385.

Sodroski JG (1992). The human T-cell leukemia virus (HTLV) transactivator (Tax) protein. *Biochem Biophys Acta* **1114**: 19–29.

Suzuki T, Fujisawa J-I, Toita M, Yoshida M (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.

Tillmann M, Wessner R, Wigdahl B (1994). Identification of human T-cell lymphotropic virus type I 21-base-pair repeat-specific and glial cell-specific DNA-protein complexes. *J Virol* **68**: 4597–4608.

Tillmann M, Krebs FC, Wessner R, Pomeroy SM, Goodenow MM, Wigdahl B (1994). Neuroglial-specific factors and the regulation of retrovirus transcription. *Adv Neuroimmunol* **4**: 305-318.

Tillmann M, Wigdahl B (1994). Identification of HTLV-I 21 bp repeat-specific DNA-protein complexes. *Leukemia* **8**: S83–S87.

Watanabe T, Seiki M, Yoshida M (1984). HTLV type I (US isolate) and ATLV (Japanese isolate) are the same species of human retrovirus. *Virology* **133**: 238–241.

Wong-Staal F, Gallo RC (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.

Zhao LJ, Giam CZ (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.