Conformation of Tax-response elements in the human T-cell leukemia virus type I promoter

Julia M Cox, Leslie S Sloan and Alanna Schepartz*

Department of Chemistry, Yale University, New Haven, CT 06520-8107, USA

Background: HTLV-I Tax is believed to activate viral gene expression by binding bZIP proteins (such as CREB) and increasing their affinities for proviral TRE target sites. Each 21 bp TRE target site contains an imperfect copy of the intrinsically bent CRE target site (the TRE core) surrounded by highly conserved flanking sequences. These flanking sequences are essential for maximal increases in DNA affinity and transactivation, but they are not, apparently, contacted by protein. Here we employ non-denaturing gel electrophoresis to evaluate TRE conformation in the presence and absence of bZIP proteins, and to explore the role of DNA conformation in viral transactivation.

Results: Our results show that the TRE-1 flanking sequences modulate the structure and modestly increase the affinity of a CREB bZIP peptide for the TRE-1 core recognition sequence. These flanking sequences are also essential for a maximal increase in stability of the CREB–DNA complex in the presence of Tax.

Conclusions: The CRE-like TRE core and the TRE flanking sequences are both essential for formation of stable CREB–TRE-1 and Tax–CREB–TRE-1 complexes. These two DNA segments may have co-evolved into a unique structure capable of recognizing Tax and a bZIP protein.

Chemistry & Biology December 1995, **2**:819–826

Key words: CREB, cyclic-AMP response element, DNA bending, Tax, transcriptional activation

Introduction

Human T-cell leukemia virus type I (HTLV-I) is the etiologic agent of a group of diseases known collectively as adult T-cell leukemia [1,2]. One of the proteins encoded by HTLV-I is the transcriptional activator, Tax, which regulates viral gene expression through three 21-base-pair (bp) DNA target sites located within the HTLV-I long terminal repeat (LTR) [3]. These 21-bp target sites are known as Tax response elements (TREs). Each TRE contains an imperfect copy of the cyclic-AMP response element (CRE, ATGACGTCAT), the target site for bZIP proteins in the CREB/ATF sub-family [4]. This CRE target site is flanked by highly conserved sequences of unknown but critical function (Fig. 1a) [5]. Although Tax activates transcription through target sites recognized by bZIP proteins, Tax is not a bZIP protein, and does not bind DNA with high affinity [6,7]. Substantial evidence indicates that Tax activates HTLV-I transcription by interacting with host bZIP proteins to increase their affinities for viral TREs [8–12]. bZIP proteins that interact with Tax and bind TREs include CREB and ATF-2, two proteins that bind the consensus CRE target site with high affinity [6,13–18] and specificity [19,20]. Considerable efforts are currently aimed at understanding the mechanism by which Tax interacts with bZIP proteins to increase their affinities for DNA [17,21–29].

Two significant differences exist between the three 21-bp TREs and a consensus CRE target site (Fig. 1). First, only six or seven bp of the central 10-bp core of each TRE are identical to the corresponding base-pairs of the CRE target site. Second, whereas the sequences flanking the 10-bp CRE target site do not always influence the DNA-binding activities or transcriptional potencies of CREB/ATF proteins [23,30,31], the sequences flanking the central core of each TRE are critical for both functions. *In vitro* selection experiments indicate that the flanking sequences contribute to the stability of the Tax–CREB–TRE complex but not to that of the CREB–TRE complex [23,25,30]. Mutagenesis experiments indicate that transcriptional activation of the HTLV-I genome is abolished when the flanking sequences are mutated [5,8,32,33].

Several possible explanations exist for the importance of the conserved TRE flanking sequences. These sequences could be contacted by Tax, or they could be contacted by CREB in the presence of Tax, or they could alter the overall conformation of the DNA so that it presents an appropriate binding site for a Tax–CREB complex. The CREB–TRE and Tax–CREB–TRE complexes produce similar interference footprints, suggesting that the flanking sequences are not contacted by Tax or by CREB when Tax is present [23,24]. Two observations suggest a role for DNA conformation in the formation of Tax–CREB–TRE complexes. First, d(C)d(G) runs, which are found in the TRE flanking sequences, exhibit small but detectable curvature in solution [34]. Second, the consensus CRE target site, which resembles the central TRE core, displays major-groove curvature that is no longer seen after binding by members of the CREB/ATF family of bZIP proteins such as CREB [20,35]. Given the possibility of non-B-form DNA structure in the TRE target site, we examined its

*Corresponding author.

 ISSN 1074-5521

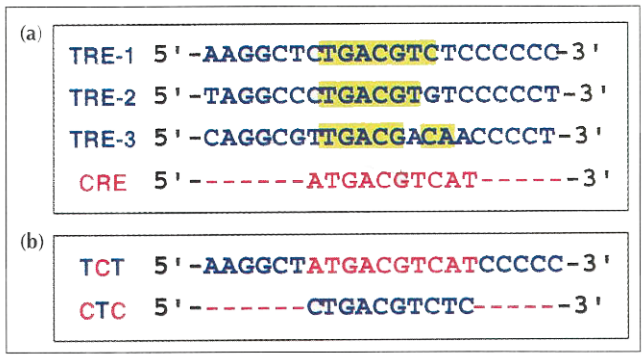

Fig. 1. Sequences of the TREs, CRE and the TCT and CTC chimeric target sites. **(a)** Sequences of the three Tax response elements (TREs) found in the HTLV-I long terminal repeat and a consensus CRE target site. Bases within each TRE target site that are identical to the corresponding CRE target site base are highlighted in yellow. **(b)** Sequences of the chimeric TCT and CTC sequences used to further examine TRE-1 conformation. Pink letters indicate bases derived from the CRE target site, and blue letters indicate bases derived from the TRE-1 target site.

conformation in the presence and absence of bZIP proteins to gain insight into the role of DNA conformation in the formation of Tax–CREB–TRE complexes.

Results and discussion

Phasing analysis of TRE-1

Helical phasing analysis [36] was employed to determine whether the 21-bp TRE-1 target site, like the CRE target site, is intrinsically bent. Helical phasing analysis is based on the differential mobilities of curved and straight DNA fragments in non-denaturing polyacrylamide gels. Bent DNA molecules migrate more slowly through these gels than do straight molecules of the same length. In a phasing analysis, a set of DNA test fragments is constructed. Each test fragment contains a different number of base pairs between the DNA target site of interest and an A-tract sequence which bends by 54° toward the minor groove [37]. The distance between the centers of these sites is varied in five steps over the space of one helical turn (10–11 bp). If the DNA target site of interest is straight, then the five test fragments will migrate through a non-denaturing gel at comparable rates because they will all contain only one bent sequence — the A-tract. If the DNA target site of interest is intrinsically bent, however, then the five fragments will migrate through the gel at different rates depending on whether the target-site bend increases or decreases the overall curvature of the DNA fragment. Phasing analysis distinguishes bends that are a result of intrinsic curvature from those that are a result of inherent flexibility, and it also determines the direction of the target bend relative to the A-tract bend [38].

The three TRE target sites in the HTLV-I LTR are not identical. We chose to study TRE-1, whose central 10-bp core is most similar to the consensus CRE target site (Fig. 1a). The test fragments used for phasing analysis of TRE-1 and CRE (as a control) are shown in Figure 2a. The CRE test fragments contain flanking sequences that do not themselves exhibit intrinsic curvature [35]. The TRE test fragments include the conserved flanking sequences found in the 21-bp TRE-1 target site (Fig. 1a), and therefore have a greater number of base pairs separating the target site and the A-tract than do the corresponding CRE test fragments. Since the TRE constructs otherwise contain the same sequences as the CRE constructs, any observed curvature in the TRE target site cannot result from an intrinsic bend in the linker [35].

An autoradiogram illustrating the relative mobilities of the TRE-1 test fragments is shown in Figure 2b; this figure also shows test fragments containing either the AP-1 or the CRE target site. The data show that while the mobilities of the five CRE test fragments vary in a phase-dependent manner, the mobilities of the five TRE-1 test fragments do not. Plots showing the relative mobilities of the two sets of test fragments as a function of the distance in base pairs between the centers of the TRE-1 or CRE target site and the center of the A-tract are shown in Figure 2c. Among the CRE test fragments, the molecule with the slowest mobility (CRE-26) contains a linker that separates the center of the CRE target site from the center of the A-tract by 2.5 helical turns of DNA. In this molecule, the CRE target site and the A-tract are out of phase, confirming the presence of an intrinsic major-groove bend in the CRE target site [35] that adds constructively to the A-tract bend in this orientation. The five TRE test fragments, however, exhibit the same mobility, indicating the absence of a net intrinsic bend in the 21-bp TRE target site. Estimation of the bend angles [39–41] suggests that the CRE target site is bent by 11 ± 1°, while the TRE-1 target site is straight (1 ± 1°). Because the TRE-1 sequence differs from the consensus CRE target site by only 3 bp in the central 10-bp DNA binding region, the finding that the TRE-1 target site has no curvature was unexpected.

Three possible explanations could account for the difference in intrinsic curvature between the TRE-1 and CRE target sites. The first possibility is that the 5–6 bp flanking either side of the 10-bp CRE-like core in the TRE-1 target site might modulate the intrinsic major-groove bend and cause TRE-1 to appear straight in the phasing analysis. This model invokes a 21-bp TRE with several composite bends that together generate a fragment that appears straight in a phasing analysis. A second possibility is that the 3-bp difference between the 10-bp core of the CRE and TRE-1 target sites might result in a change in conformation. This 3-bp difference actually leads to five new dinucleotide steps that could be important [42]. Finally, it is possible that a combination of both the conserved TRE flanking sequences and the central base pair changes could lead to the observed TRE-1 conformation.

Phasing analysis of CRE/TRE chimeras

To distinguish between these possibilities, we constructed two sets of phasing oligonucleotides, labeled TCT and CTC in Figures 1b and 2a. The TCT test fragments contain the 10-bp consensus CRE target site

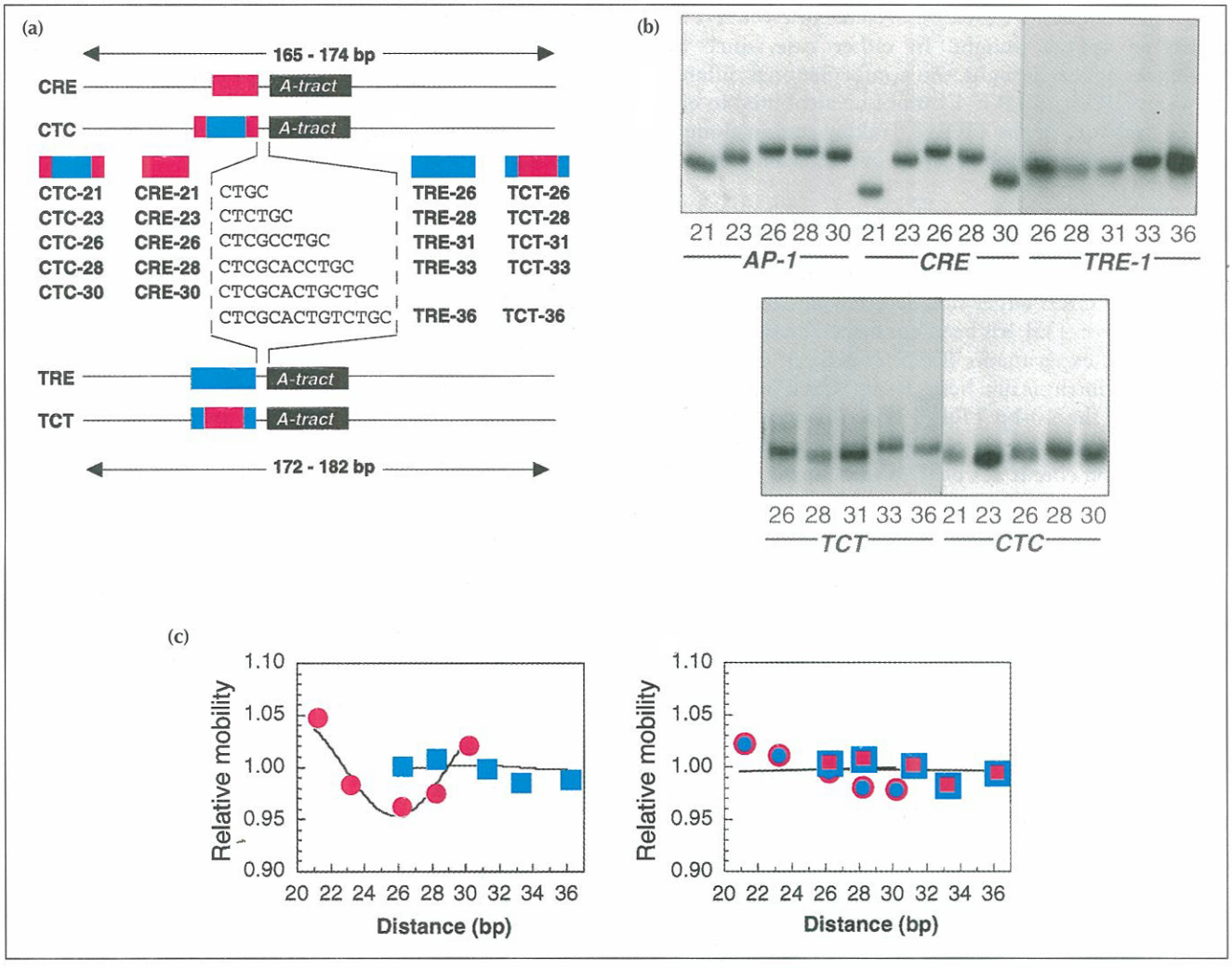

Fig. 2. Phasing analysis of DNA curvature demonstrates that the CRE target site is bent, but the TRE-1, CTC and TCT target sites are not. **(a)** The probes used for phasing analysis of the CRE and TRE-1 target sites contain a CRE (pink boxes) or TRE-1 (blue boxes) target site separated by a variable length linker from a 25-bp A-tract sequence (black boxes) that bends intrinsically toward the minor groove by 54° [37]. The TRE test fragments include the conserved flanking sequences found in the 21-bp TRE target site and therefore have a greater number of base pairs separating the TRE target site and the A-tract. The constructs used for phasing analysis of the chimeric CTC (blue box surrounded by pink) and TCT (pink box surrounded by blue) sequences were constructed similarly. The probe names refer to the distance in base pairs between the centers of the test target site and the 25-bp A-tract. The probes containing the TRE flanking regions (TRE and TCT) have a larger number of base pairs in this region due to the increased length of the test target site. With the exception of the variable linker, all five constructs were the same size and contained the same sequence. **(b)** Autoradiogram illustrating non-denaturing electrophoretic mobility analysis of the AP-1, CRE, TRE-1, TCT, and CTC DNA test fragments (numbering as in (a)). **(c)** Plots of the relative mobilities of the CRE (pink), TRE-1 (blue), CTC (pink and blue circles), and TCT (blue and pink squares) DNA test fragments as a function of the distance in base pairs between the centers of the target and the A-tract sites. The mobilities of the DNA fragments (μ_f) were measured as the distance in millimeters from the center of the electrophoresis well to the center of the DNA band and were normalized to the average mobility of the fastest and slowest fragments to give the relative mobilities ($\mu_{f,n} = \mu_f/\mu_{f(avg.)}$). The data shown represent the average of at least three separate determinations. Error bars represent one standard deviation. The points are connected by the calculated best fit of the data to a cosine function [41].

surrounded by the TRE-1 flanking sequences; the CTC test fragments contain the 10-bp CRE-like core of TRE-1 surrounded by the CRE flanking sequences [35]. If the TRE-1 flanking sequences counteract a bend in the central 10-bp core, then the TCT test fragments should not show evidence of intrinsic curvature while the CTC test fragments should. If the three-base-pair difference between the CRE target site and the TRE-1 central core abolishes the intrinsic curvature, then the TCT test fragments should show evidence of intrinsic curvature and the CTC test fragments should not. If both factors contribute to the conformation of the DNA, then both sets of chimeric test fragments should appear either straight or bent.

Figure 2b shows the phasing analysis of the TCT and CTC test fragments. Neither set shows evidence of intrinsic curvature. Relative mobility plots (Fig. 2c) show no differences among the sets of CTC and TCT oligonucleotides. The bend angles estimated for both chimeras were found to be 0 ± 1° [39–41]. Either the two sets of chimeric DNAs are straight, or they contain composite

bends that counteract each other and cause the test fragments to appear straight. In either case, since both sequences appear straight, the conformational difference between CRE and TRE-1 cannot be attributed to either the central core or the TRE-1 flanking region alone, but must be an essential combination of these elements.

The TCT and CTC phasing data can be analyzed in light of other results to provide more specific conclusions about the nature of the TRE and CRE target sites. The intrinsic bend in the CRE target site, detected initially by helical phasing analysis [35], has been confirmed by the use of ligation-ladder experiments (L.S.S. & A.S., in preparation). When the intrinsically bent 10-bp CRE target site is replaced by the 10-bp TRE-1 core (to create the CTC construct), no bending is seen. This result suggests that the TRE-1 10-bp core is not bent, in agreement with ligation ladder analysis of multimers of the 10-bp TRE-1 central core (J.M.C. & A.S., unpublished results). It also suggests that a 3-bp (5 base-step) change is enough to eliminate the intrinsic CRE bend, which is consistent with mutagenesis data showing that the CRE bend requires two out-of-phase 5'-TGA-3' half sites located on opposite DNA strands (L.S.S. & A.S., in preparation). This latter suggestion is also consistent with the finding of major-groove curvature within protein-bound DNA segments containing two pyrimidine-purine (YR) steps placed 6–10 bp apart with an additional YR step in the center [42]. This arrangement is found in the CRE target site but not in the TRE-1 10-bp core.

When the intrinsically bent 10-bp CRE target site is surrounded by the TRE flanking sequences in the TCT constructs, the CRE major-groove bend is lost. This result suggests that the TRE-1 target site flanking sequences are sufficient to alter DNA conformation at a distance, and that perhaps the central 10-bp core and the conserved flanking sequences somehow act together to provide the unique structure of the 21-bp TRE target site.

Effect of CREB on DNA conformation

Despite the differences in their conformations, the TRE and CRE target sites are bound by some of the same bZIP proteins, such as CREB. CREB is one of several bZIP proteins that removes the intrinsic CRE target site bend when it binds [20,35]. Therefore, it is possible that the TRE-1 conformation has evolved to stabilize CREB–TRE complexes selectively. The TRE-1 target

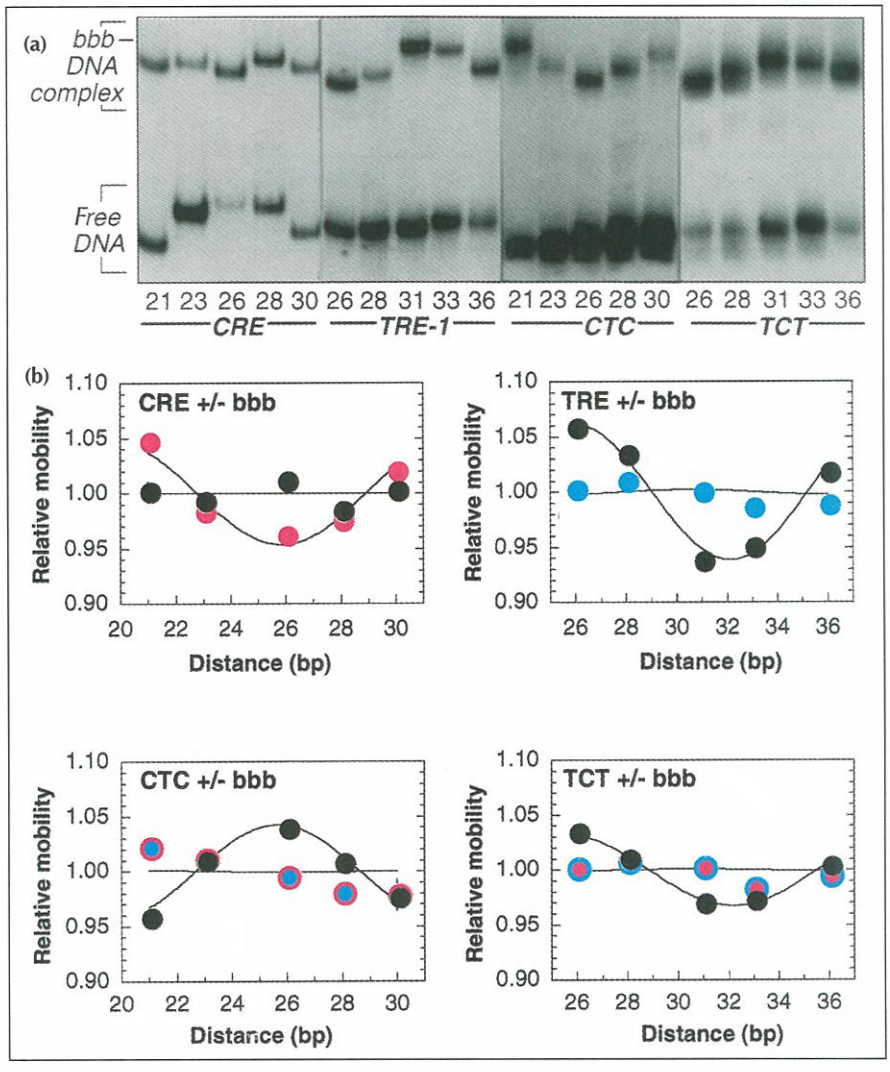

Fig. 3. Binding of a CREB bZIP peptide removes the curvature of the CRE target site, but induces bends in the TRE-1, CTC and TCT target sites. **(a)** Autoradiogram illustrating an electrophoretic mobility shift analysis [46,47] of the CREB bZIP element peptide bbb [20, 45] bound to the CRE, TRE-1, CTC and TCT phasing oligonucleotides. **(b)** Plots of the relative mobilities ($\mu_{f,n}$, defined in Fig. 2 legend) of the free CRE (pink), TRE (blue), CTC (blue in pink), and TCT (pink in blue) test fragments and the corresponding bbb-bound complexes (black symbols) as a function of the distance in base pairs between the centers of the target and the A-tract sites. Plots were created as described in Figure 2. Error bars represent one standard deviation.

site could, for example, present a 'pre-straightened' CRE target site that is organized for the binding of CREB. Alternatively, the TRE-1 target site could behave like the CRE target site and bend to present an appropriate conformation for CREB. To explore the relevance of the different conformations for protein binding, we evaluated the conformation of each set of phasing oligonucleotides (CRE, TRE-1, CTC, and TCT) bound to a peptide comprising the bZIP element of human CREB, bbb [20].

An autoradiogram illustrating the mobilities of the free and bbb-bound CRE, TRE-1, CTC and TCT test fragments is shown in Figure 3a. The data show that bbb alters DNA conformation in all four cases. The bbb-bound CRE test fragments have lost the phase-dependent mobility differences seen in the free DNAs [20], whereas the bbb-bound TRE-1, TCT and CTC test fragments have gained phase-dependent mobility differences that are absent in the free DNAs. Figure 3b shows the relative mobilities of the free and bound DNA for each set of phasing constructs plotted against the distance between the centers of the target site and the A-tract. Since the lowest mobility complexes result when the centers of the test and A-tract sites are in phase, we conclude that the bbb–TRE-1, bbb–TCT and bbb–CTC complexes contain DNA bent toward the minor groove. The observation that formation of the bbb–CRE and bbb–TRE-1 complexes both involve a bend toward the minor groove (11 ± 1° and 14 ± 1°, respectively) indicates that the TRE target site does not present a straight CRE target site that is pre-organized for binding CREB. The CTC and TCT target sites also bend toward the minor groove upon binding of bbb; CTC bends by 9 ± 1° and TCT by 7 ± 1°. The observation that changes in flanking DNA sequence alter the extent of bending induced by bbb supports the idea that the central 10-bp core and the conserved flanking sequences act together to provide the unique structure of the 21-bp TRE target site.

Effect of flanking sequences on Tax–bZIP–DNA stability

To determine if a correlation exists between TRE/CRE conformation and the extent to which bZIP peptide binding to these sequences can be enhanced by Tax, we measured the effect of added Tax on the stabilities of the CRE, TRE, TCT, and CTC complexes of bbb (Fig. 4). The bbb–TRE complex exhibits an equilibrium dissociation constant (K_d) of 4.5×10^{-16} M^2 (ΔG_{obs} = −19.4 kcal mol^{-1}) in the absence of Tax and 4.4×10^{-17} M^2 (ΔG_{obs} = −20.7 kcal mol^{-1}) in the presence of Tax, corresponding to a 1.3 kcal mol^{-1} increase in binding free energy due to Tax ($\Delta\Delta G_{Tax}$). The bbb–CTC complex exhibits a K_d of 8.1×10^{-16} M^2 (ΔG_{obs} = −19.1 kcal mol^{-1}) in the absence of Tax and 3.0×10^{-16} M^2 (ΔG_{obs} = −19.7 kcal mol^{-1}) in the presence of Tax, corresponding to a 0.6 kcal mol^{-1} increase in binding free energy due to Tax. The bbb–CRE complex exhibits a K_d of 8.8×10^{-16} M^2 (ΔG_{obs} = −19.1 kcal mol^{-1}) in the absence of Tax and 1.7×10^{-16} M^2 (ΔG_{obs} = −20.0 kcal mol^{-1}) in the presence of Tax, corresponding to a 0.9 kcal mol^{-1} increase in binding free energy due to Tax. Finally, the bbb–TCT complex exhibits a K_d of 8.7×10^{-16} M^2 (ΔG_{obs} = −19.1 kcal mol^{-1}) in the absence of Tax and 1.7×10^{-16} M^2 (ΔG_{obs} = −20.0 kcal mol^{-1}) in the presence of Tax, corresponding to a 0.9 kcal mol^{-1} increase

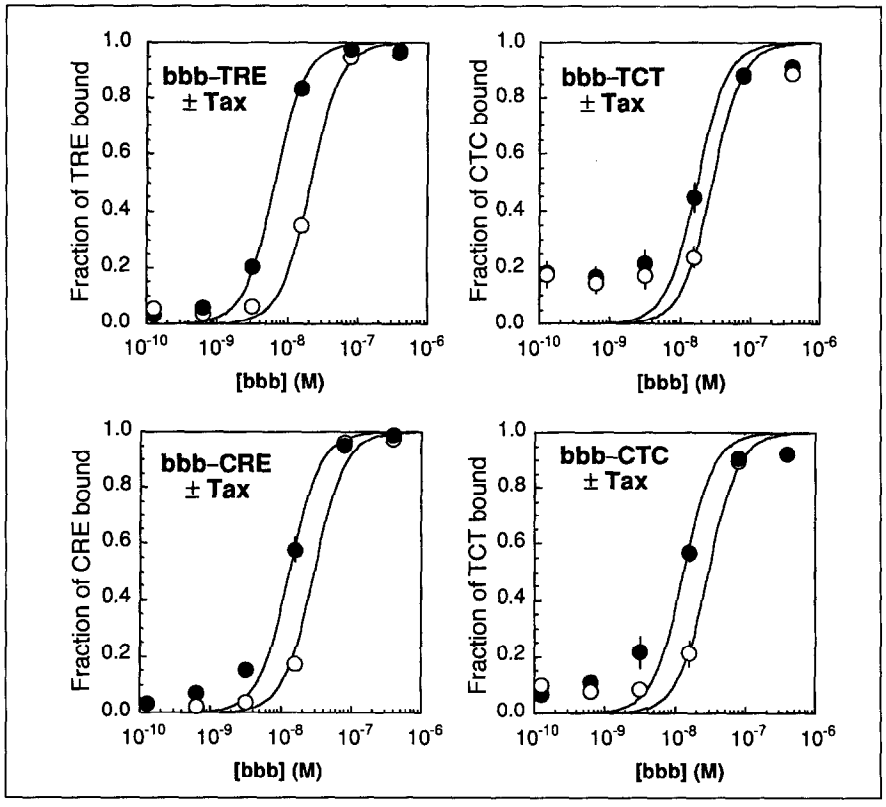

Fig. 4. Determination of equilibrium dissociation constants of bbb–DNA complexes in the presence and absence of Tax. Quantitative electrophoretic mobility shift analyses were performed to determine the stabilities of bbb–DNA complexes in the presence (●) and absence (○) of Tax. The data shown represent the average of three separate determinations. Error bars represent one standard deviation. Plots were created as described in Materials and methods.

in binding free energy due to Tax. Thus, the effect of Tax on stability of the bZIP–DNA complex is largest (1.3 kcal mol^{-1}) within the context of a natural TRE-1 target site. Changes in either the flanking sequences or the sequence of the central core or both diminish the effect of Tax to 0.9 or 0.6 kcal mol^{-1}. These data suggest that the extent to which a bZIP peptide is responsive to Tax correlates directly with the detailed structure of the DNA target site to which the bZIP peptide binds. This observation provides additional support for a model in which Tax recognizes the detailed structure of a bZIP–DNA complex, and not simply bZIP peptide sequences [21].

Implications for Tax recognition of CREB–TRE complexes

Our data show that the TRE-1 target site must bend upon formation of a CREB–TRE-1 complex, indicating that the TRE-1 target site structure is not pre-organized for CREB. The TRE-1 target site could, however, be pre-organized for binding Tax and CREB. Tax could bend the TRE target site to aid CREB binding, or it could pre-organize CREB such that it binds the TRE-1 target site without inducing the DNA bending seen with CREB alone. Both scenarios could explain the stabilization of the CREB–DNA complex. Because the Tax–CREB–TRE complexes are unstable during the native gel electrophoresis conditions of a phasing analysis, we could not, in this study, determine the effect of Tax on the conformation of the CREB–CRE and CREB–TRE complexes. Further work is in progress to address this issue.

Significance

HTLV-I is one of several viral pathogens whose replication depends on a viral protein that expands or alters the DNA recognition properties of a cellular transcription factor. Other examples include hepatitis B virus, which produces the pX protein, bovine leukemia virus, which produces BLV Tax, and adenovirus, which produces the E2 protein. In the case of HTLV-I Tax, the DNA sequences recognized by the human bZIP protein CREB are expanded from 8–10 to 21 bp, with the additional nucleotides flanking a non-consensus bZIP target site. Previous work has suggested that these flanking sequences are not contacted by CREB or Tax, yet they are essential for the observed changes in DNA-binding and viral transactivation. Here we provide evidence that the flanking sequences are conserved to create, through effects on DNA conformation, a better binding site for the complexes formed between Tax and the cellular bZIP transcription factors that are required for HTLV-I replication.

Materials and methods

Preparation of test fragments for TRE-1 phasing analysis

The single-stranded oligonucleotides required to generate the TRE-1 test fragment set were synthesized on a Millipore Expedite model 8909 nucleic-acid synthesizer. These oligonucleotides contained the following sequences: TRE-26S-A: 5'-CTAGAAGGCTCTGACGTCTCCCCCC, TRE-28S-A: 5'-CTAGAAGGCTCTGACGTCTCCCCCCCT, TRE-31S-A: 5'-CTAGAAGGCTCTGACGTCTCCCCCCCCTCGC, TRE-33S-A: 5'-CTAGAAGGCTCTGACGTCTCCCCCC-CTCGCAC, TRE-36S-A: 5'-CTAGAAGGCTCTGACGT-CTCCCCCCCTCGCACTGT, A-tractS-A: 5'-CTGCAAAA-CGGGCAAAAACGGGCAAAAACGC, TRE-26S-B: 5'-GCA-GGGGGGGAGACGTCAGAGCCTT, TRE-28S-B: 5'-GC-AGAGGGGGGGAGACGTCAGAGCCTT, TRE-31S-B: 5'-GCAGGCGAGGGGGGGAGACGTCAGAGCCTT, TRE-33S-B: 5'-GCAGGTGCGAGGGGGGGAGACGTCA-GAGCCTT, TRE-36S-B: GCAGACAGTGCGAGGGGG-GGAGACGTCAGAGCCTT, A-tractS-B: 5'-TCGAGCGTT-TTTGCCCGTTTTTGCCCGTTTTT, and were purified by standard methods [43]. The five TRE-S-B strands and the A-tractS-A strand were phosphorylated using T4 polynucleotide kinase, and each set of complementary A and B strands were annealed. Each of the five double-stranded TRE fragments prepared this way (TRE-26D, TRE-28D, TRE-31D, TRE-33D, TRE-36D) was joined separately to the double-stranded A-tract fragment using T4 DNA ligase in buffer containing 50 mM Tris HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP and 50 μg ml^{-1} BSA at 16 °C for 12 h. These five inserts were then cloned separately into pBEND2 [44] between the unique *Sal* I and *Xba* I restriction sites, using the ligation conditions above, to generate five pJMC-TRE plasmids. These plasmids were used to transform competent *Escherichia coli* GM2929 cells and plasmid DNA was isolated from individual colonies using the Wizard Miniprep Kit (Promega). Correct clones were identified initially by loss of the Sal I site. Plasmids were sequenced (Circumvent Kit, New England Biolabs) to ensure the integrity of the inserted sequences. Digestion of each plasmid with *Stu* I generated the set of five oligonucleotides used for phasing analysis. These phasing oligonucleotides were amplified and internally labeled with α-[^{32}P]-ATP using the PCR and Vent (exo-) DNA polymerase (New England Biolabs) in buffer containing 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 6 mM $MgSO_4$, 0.1 % Triton X-100, 0.1 μg ml^{-1} single stranded binding protein (Promega), 0.1 mg ml^{-1} BSA.

Preparation of test fragments for TCT and CTC phasing analysis

Six oligonucleotides were synthesized to make each set of phasing oligonucleotides. The sequences used for the TCT oligonucleotides were: TCT Main: 5'-GGACTAGTCTCGAGTT-TAAAGATATCCAGCTGCCCGGGAGGCCTTCGC-GAAATATTGGTACCCCATGGAATCGAGGGATCCTC-TAGAAGGCTATGACGT, TCT-26S: 5'-ATGGATCGATGC-TAGCAGATCTACGCGTGTCGAGCGTTTTTGCC-CGTTTTTGCCCGTTTTTGCAGGGGGGATGA-CGT-CATAGCCTTCTAGAGGATC, TCT-28S: 5'-ATGGATCG-ATGCTAGCAGATCTACGCGTGTCGAGCGTTTT-TGCCCGTTTTTGCCCGTTTTTGCAGAGGGGGGAT-GACGTCATAGCCTTCTAGAGGATC; TCT-31S: 5'-ATG-GATCGATGCTAGCAGATCTACGCGTGTCGAG-CGTTTTTGCCCGTTTTTGCCCGTTTTTGCAGGC-GAGGGGGGATGACGTCATAGCCTTCTAGAGGATC, TCT-33S: 5'-ATGGATCGATGCTAGCAGATCTACGCG-TGTCGAGCGTTTTTGCCCGTTTTTGCCCGTTT-TTGCAGGTGCGAGGGGGGATGACGTCATAGCCTTC-TAGAGGATC, TCT-36S: 5'-ATGGATCGATGCTAGCA-GATCTACGCGTGTCGAGCGTTTTTGCCCGTT-TTTGCCCGTTTTTGCAGACAGTGCGAGGGGGGAT-

GACGTCATAGCCTTCTAGAGGATC. The sequences used for the CTC oligonucleotides were: CTC-Main: 5'-ACTAGTCTCGAGTTTAAAGATATCCAGCTGCCCGGGAGGCCTTCGCGAAATATTGGTACCCCATGGAATCGAGGGATCCTCTAGAGAGCTGA, CTC-21S: 5'-CCATGGATCGATGCTAGCAGATCTACGCGTGTCGAGCGTTTTTGCCCCGTTTTTTGCCCCGTTTTTTGCAGGAGACGTCAGCTCTCTAGAGGATCCCT, CTC-23S: 5'-CCATGGATCGATGCTAGCAGATCTACGCGTGTCGAGCGTTTTTTGCCCCGTTTTTTGCCCCGTTTTTTGCAGAGGAGACGTCAGCTCTCTAGAGGATCCCT, CTC-26S: 5'-CCATGGATCGATGCTAGCAGATCTACGCGTGTCGAGCGTTTTTTGCCCCGTTTTTTGCCCCGTTTTTTGCAGGCGAGGAGACGTCAGCTCTCTAGAGGATCCCT, CTC-28S: 5'-CCATGGATCGATGCTAGCAGATCTACGCGTGTCGAGCGTTTTTTGCCCCGTTTTTTGCCCCGTTTTTGCAGGTGCGAGGAGACGTCAGCTCTCTAGAGGATCCCT, CTC-30S: 5'-CCATGGATCGATGCTAGCAGATCTACGCGTGTCGAGCGTTTTTTGCCCCGTTTTTTGCCCCGTTTTTTGCAGCAGTGCGAGGAGACGTCAGCTCTCTAGAGGATCCCT. Mutually primed synthesis experiments performed with TCT-Main or CTC-Main and each of their five respective TCT and CTC oligonucleotides created phasing constructs identical to TRE-1 and CRE test fragments with their central 10-bp cores switched, as shown in Figure 1b. Oligonucleotides were purified on an 8 % denaturing polyacrylamide gel, eluted from the gel by effusion into TE buffer (10 mM Tris HCl, 1 mM EDTA (pH 8.0)), and desalted on Sephadex G25 columns. Concentrations were determined by measuring absorbance at 260 nm using the relation 1 OD = 33 $\mu g\ ml^{-1}$. Mutually primed synthesis reactions using Vent (exo-) DNA polymerase (New England Biolabs) were used to make each of the ten double-stranded phasing oligonucleotides in buffer containing 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 6 mM MgSO4, 0.1 % Triton X-100, 0.1 $\mu g\ ml^{-1}$ single-stranded binding protein (Promega), 0.05 mg ml^{-1} BSA. DNA was internally labeled with α-[^{32}P]-ATP during the extension. Extension products were phenol/chloroform extracted and ethanol precipitated, then purified on a 10 % denaturing polyacrylamide gel and re-annealed prior to phasing analysis experiments.

Phasing analysis

Phasing analyses in the absence of bbb were performed using 8 % non-denaturing (32:1 (wt/wt) acrylamide/N,N'-methylene bisacrylamide) polyacrylamide gels prepared in TG buffer (19 mM Tris base, 162 mM glycine (pH 8.9)) and run at 4 °C for 1500 V h. Phasing analyses performed in the presence of bbb were carried out as described above, except that the DNA test fragments (10 to 100 pM) were incubated with bbb (30–100 nM) in phosphate-buffered saline (2.7 mM KCl, 137 mM NaCl, 4.3 mM Na_2HPO_4 and 1.4 mM KH_2PO_4 (pH 7.3)) containing 5 % glycerol, 1 mM EDTA, 1 mM dithiothreitol, 0.8 mg ml^{-1} BSA and 0.05 % NP-40 for 30 min at 4 °C prior to electrophoresis. Radioactivity was quantified with a Betascope 603 Blot Analyzer (Betagen Corporation, Waltham MA) and by autoradiography.

Protein expression

Expression of bbb was carried out in Bl21(DE3)pLys S *E. coli* from expression vector $pCREB_{259\text{-}327}$ [45] and the protein was purified to homogeneity by hydroxyapatite chromatography and HPLC as described [20]. Tax was expressed in MC1061 *E. coli* from the expression vector pTaxH6 [12] and was purified to homgeneity by ammonium sulfate precipitation and Ni-chelate chromatography.

Electrophoretic mobility-shift experiments

Binding reactions contained the indicated peptide and <50 pM of a 33-bp end-labeled DNA probe (TRE: CTCTGCAAGGCTCTGACGTCTCCCCCCCTCTGC, CRE: CTCTGCGTGGAGATGACGTCATCTCGTCTCTGC, TCT: CTCTGCAAGGCTATGACGTCATCCCCCCTCTGC, CTC: CTCTGCGTGGAGCTGACGTCTCCTCGTCTCTGC) in a final reaction mixture containing 65 mM HEPES (pH 7.1), 15 mM HEPES (pH 7.9), 90 mM KCl, 7.5 mM $MgCl_2$, 6 μM $ZnSO_4$, 600 μM EDTA, 100 $\mu g\ ml^{-1}$ BSA, 6 mM β-mercaptoethanol, 0.07 % NP-40, 14 % glycerol, and approximately 0.5 μM Tax (when present) in a final volume of 10 μl. Binding reactions were incubated at 4 °C for 30 min and then applied to a running 16 x 18 cm non-denaturing 10 % polyacrylamide (49:1 acrylamide:bisacrylamide) gel prepared in 153 mM glycine, 20 mM Tris base (pH 8.5). Samples were loaded and subjected to electrophoresis at 500 V for 45–60 min. Gels were maintained at 4 °C during electrophoresis by immersion in a circulating, temperature-controlled water bath. The amounts of complexed and free DNA were quantified on a Betagen 603 Blot Analyzer (Betagen Inc., Waltham, MA). Equilibrium dissociation constants of peptide–CRE complexes were obtained by fitting the binding data to the Langmuir equation

$$\Theta = \frac{1}{1 + \frac{K_d}{([\text{peptide}]_T)^2}}$$

where K_d is the adjustable parameter and Θ = fraction of DNA bound.

Values for ΔG_{obs} were calculated from the relationship $\Delta G_{obs} = -RT\ \ln(1/K_d)$ where R = 0.00198 kcal $mol^{-1}\ K^{-1}$ and T = 277 K. Values for $\Delta\Delta G_{Tax}$ are valid with the assumption that Tax does not alter the standard state, as defined by the binding conditions in the absence of Tax.

Acknowledgements: We are grateful to G. Perini and M.R. Green for sharing their $TaxH_6$ purification procedure with us. This work was supported by the NIH and the National Foundation for Cancer Research. L.S.S. is an NSF predoctoral fellow. A.S. is a David and Lucile Packard Fellow, an NSF Presidential Young Investigator, a Camille and Henry Dreyfus Teacher-Scholar, an Alfred P. Sloan Foundation Research Fellow, and a 1995 Arthur C. Cope Scholar.

References

1. Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. & Gallo, R.C. (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. USA* **77**, 7415–7419.
2. Yoshida, M., Miyoshi, I. & Hinuma, Y. (1982). Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc. Natl. Acad. Sci. USA* **79**, 2031–2035.
3. 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.
4. Hurst, H.C. (1994). Transcription factors 1: bZIP proteins. *Protein Profile* **1**, 123–168.
5. Jeang, K.-T., Boros, I., Brady, J., Radonovich, M. & Khoury, G. (1988). Characterization of cellular factors that interact with the human T-cell leukemia virus type I p40x-responsive 21-base-pair sequence. *J. Virol.* **62**, 4499–4509.

6. 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.
7. Beimling, P. & Moeling, K. (1988). Isolation and characterization of the Tax protein of HTLV-I. *Oncogene* **4**, 511–516.
8. Park, R.E., Haseltine, W.A. & Rosen, C.A. (1988). A nuclear factor is required for transactivation of HTLV-I gene expression. *Oncogene* **3**, 275–279.
9. Nyborg, J.K., Dynan, W.S., Cjem, I.S.Y. & Wachsman, W. (1988). Binding of host-cell factors to DNA sequences in the long terminal repeat of human T-cell leukemia virus type I: implications for viral gene expression. *Proc. Natl. Acad. Sci. USA* **85**, 1457–1461.
10. 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.
11. Matthews, M.-A.H., Markowitz, R.-B. & Dynan, W.S. (1992). *In vitro* activation of transcription by the human T-cell leukemia virus type I Tax protein. *Mol. Cell Biol.* **12**, 1986–1996.
12. Zhao, L.J. & Giam, C.Z. (1991). Interaction of the human T-cell lymphotrophic virus type I (HTLV-I) transcriptional activator Tax with cellular factors that bind specifically to the 21-base-pair repeats in the HTLV-I enhancer. *Proc. Natl. Acad. Sci. USA* **88**, 11445–11449.
13. Beimling, P. & Moelling, K. (1992). Direct interaction of CREB protein with 21 bp Tax-response elements of HTLV-I LTR. *Oncogene* **7**, 257–262.
14. Zhao, L.J. & Giam, C.Z. (1992). Human T-cell lymphotropic virus type I (HTLV-I) transcriptional activator, Tax, enhances CREB binding to HTLV-I 21-base-pair repeats by protein–protein interaction. *Proc. Natl. Acad. Sci. USA* **89**, 7070–7074.
15. Franklin, A.A., *et al.*, & Nyborg, J.K. (1993). Transactivation by the human T-cell leukemia virus Tax protein is mediated through enhanced binding of activating transcription factor-2 (ATF-2) ATF-2 response and cAMP element-binding protein (CREB). *J. Biol. Chem.* **268**, 21225–21231.
16. Suzuki, T., Fujisawa, J.-i., Toita, M. & Yoshida, M. (1993). The transactivator 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.
17. Wagner, S. & Green, M.R. (1993). HTLV-I Tax protein stimulation of DNA binding of bZIP proteins by enhancing dimerization. *Science* **262**, 395–399.
18. Yoshimura, T., Fujisawa, J.-i. & Yoshida, M. (1990). Multiple cDNA clones encoding nuclear proteins that bind to the Tax-dependent enhancer of HTLV-I: all contain a leucine zipper structure and basic amino acid domain. *EMBO J.* **9**, 2537–2542.
19. Metallo, S.J. & Schepartz, A. (1994). Distribution of labor among CRE-BP1 bZIP segments in the control of DNA affinity and specificity. *Chemistry & Biology* **1**, 143–151.
20. Hamm, M.K. & Schepartz, A. (1995). Studies on the formation of DNA–protein interfaces: DNA specificity and straightening by CREB. *Bioorg. Med. Chem. Lett.* **5**, 1621–1626.
21. Baranger, A.M., *et al.*, & Schepartz, A. (1995). Mechanism of DNA binding enhancement by the HTLV-I Tax protein. *Nature* **376**, 606–608.
22. Perini, G., Wagner, S. & Green, M.R. (1995). Recognition of bZIP proteins by the human T-cell leukaemia virus transactivator Tax. *Nature* **376**, 602–605.
23. Paca-Uccaralertkun, S., *et al.*, & Giam, C.Z. (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.
24. Adya, N., Zhao, L.-J., Huang, W., Boros, I. & Giam, C.-Z. (1994). Expansion of CREB's DNA recognition specificity by Tax results from interaction with Ala-Ala-Arg at positions 282–284 near the conserved DNA binding domain of CREB. *Proc. Natl. Acad. Sci. USA* **91**, 5642–5646.
25. Anderson, M.G. & Dynan, W.S. (1994). Quantitative studies of the effect of HTLV-I Tax protein on CREB protein–DNA binding. *Nucleic Acids Res.* **22**, 3194–3201.
26. Yin, M., Paulssen, E.J., Seeler, J. & Gaynor, R.B. (1995). Protein domains involved in both *in vivo* and *in vitro* interactions between human T-cell leukemia virus type I Tax and CREB. *J. Virol.* **69**, 3420–3432.
27. Adya, N. & Giam, C.-Z. (1995). Distinct regions in human T-cell lymphotropic virus type I Tax mediate interactions with activator protein CREB and basal transcription factors. *J. Virol.* **69**, 1834–1841.
28. Goren, I., Semmes, O.J., Jeang, K.-T. & Moelling, K. (1995). The amino terminus of Tax is required for interaction with the cyclic-AMP response element binding protein. *J. Virol.* **69**, 5806–5811.
29. Brauweiler, A., Garl, P., Franklin, A.A., Giebler, H.A. & Nyborg, J.K. (1995). A molecular mechanism for human T-cell leukemia virus latency and Tax transactivation. *J. Biol. Chem.* **270**, 12814–12822.
30. Benbrook, D.M. & Jones, N.C. (1994). Different binding specificities and transactivation of variant CRE's by CREB complexes. *Nucleic Acids Res.* **22**, 1463–1469.
31. Deutsch, P.J., Hoeffler, J.P., Jameson, J.L., Lin, J.C. & Habener, J.F. (1988). Structural determinants for transcriptional activation by cAMP-responsive DNA elements. *J. Biol. Chem.* **263**, 18466–18472.
32. Montagne, J., *et al.*, & Jalinot, P. (1990). Tax1 induction of the HTLV-I 21-bp enhancer requires cooperation between two cellular DNA-binding proteins. *EMBO J.* **9**, 957–960.
33. Fujisawa, J.I., Toita, M. & Yoshida, M. (1989). A unique enhancer element for the *trans* activator ($p40^{tax}$) of human T-cell leukemia virus type I that is distinct from cyclic AMP- and 12-*O*-tetradecanoylphorbol-13-acetate-responsive elements. *J. Virol.* **63**, 3234–3239.
34. Biburger, M., Niederweis, M. & Wolfgang, H. (1994). Oligo (dCdG) runs exhibit a helical repeat of 11.1 bp in solution and cause slight DNA curvature when properly phased. *Nucleic Acids Res.* **22**, 1562–1566.
35. Paolella, D.N., Palmer, C.R. & Schepartz, A. (1994). DNA targets for certain bZIP proteins distinguished by an intrinsic bend. *Science* **264**, 1130–1133.
36. Zinkel, S.S. & Crothers, D.M. (1987). DNA bend direction by phase sensitive detection. *Nature* **328**, 178–181.
37. Koo, H.-S., Drak, J., Rice, A.J. & Crothers, D.M. (1990). Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry* **29**, 4227–4234.
38. Salvo, J.J. & Grindley, D.F. (1987). Helical phasing between DNA bends and the determination of bend direction. *Nucleic Acids Res.* **15**, 9771–9779.
39. Kerppola, T.K. & Curran, T. (1991). Fos–Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. *Cell* **66**, 317–326.
40. Kerppola, T.K. & Curran, T. (1991). DNA bending by Fos and Jun: the flexible hinge model. *Science* **254**, 1210–1214.
41. Kerppola, T.K. & Curran, T. (1993). Selective DNA bending by a variety of bZIP proteins. *Mol. Cell. Biol.* **13**, 5479–5489.
42. Suzuki, M. & Yagi, N. (1995). Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res.* **23**, 2083–2091.
43. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1987). *Molecular Cloning: A Laboratory Manual.* (2nd ed.), Cold Spring Harbor Press.
44. Kim, J., Zwieb, C., Wu, C. & Adhya, S. (1989). Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. *Gene* **85**, 15–23.
45. Santiago-Rivera, Z.I., Williams, J.S., Gorenstein, D.G. & Andrisani, O.M. (1993). Bacterial expression and characterization of the CREB bZip module: circular dichroism and 2D-^{1}H NMR studies. *Protein Sci.* **2**, 1461–1471.
46. Fried, M.G. & Crothers, D.M. (1984). Kinetics and mechanism in the reaction of gene regulatory proteins with DNA. *J. Mol. Biol.* **172**, 263–282.
47. Garner, M.M. & Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia Coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**, 3047–3060.

Received: **22 Sep 1995**; Revisions requested: **13 Oct 1995**; Revisions received: **20 Nov 1995**. Accepted: **22 Nov 1995**.