Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 5, No. 15, pp. 1621-1626, 1995
Copyright © 1995 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0960-894X/95 $9.50+0.00

0960-894X(95)00266-9

STUDIES ON THE FORMATION OF DNA•PROTEIN INTERFACES: DNA SPECIFICITY AND STRAIGHTENING BY CREB

Mary Kay Hamm & Alanna Schepartz*
Department of Chemistry
Yale University
New Haven, Connecticut 06511

Abstract: Transcriptional regulation by bZIP proteins in the CREB/ATF family requires that these proteins discriminate between the CRE and AP-1 target sites, two DNA sequences that differ by a single G•C base pair. We reported recently that CRE-BP1, one member of the CREB/ATF family, prefers the CRE target site to the AP-1 target site by 2.3 kcal•mol^{-1} and straightens the intrinsic bend in its specific CRE target site upon binding. Of all well-characterized members of the CREB/ATF family, the cyclic-AMP response-element binding protein (CREB) is among the most divergent from CRE-BP1 in primary sequence. Here we show that CREB, like CRE-BP1, (1) displays high selectivity ($\Delta\Delta G^{o}_{obs}$= 1.5 kcal•mol^{-1}) for the CRE target site and (2) straightens the intrinsically bent CRE target site upon binding. Although the effect of CREB on DNA conformation is comparable to that of CRE-BP1, the relative affinities of these two proteins for the CRE and AP-1 target sites are not. These results indicate that CRE/AP-1 specificity, the hallmark of the CREB/ATF transcription factor family, cannot be explained solely on the basis of differential DNA flexibility.

bZIP proteins[1] contain the simplest element of protein structure able to recognize DNA sequence-specifically.[2,3] The bZIP DNA-binding region (the bZIP element) is composed of fewer than sixty amino acid residues[4] (Figure 1). Each bZIP element contains a basic segment responsible for DNA binding connected through a six residue spacer to a zipper segment responsible for protein dimerization.[1] X-ray crystallographic data from three bZIP•DNA complexes show a pair of uninterrupted α–helices that interact with each other to form a parallel coiled coil and then diverge to interact with the DNA major groove.[2,3,5] In spite of the simplicity of their DNA binding regions, bZIP proteins recognize a diverse set of inverted half sites within a 9-10 base-pair sequence, a length of DNA no smaller than that recognized by proteins containing more complex DNA binding domains.[6] In addition to their ability to discriminate between target sites that differ in half-site sequence, bZIP proteins discriminate between target sites of identical half-site sequence but different half-site spacing. bZIP proteins in the AP-1 family prefer a nine base-pair AP-1 target site (ATGACTCAT) comprised of two ATGA half sites arranged in an inverted pair and separated by a single dC•dG base pair. bZIP proteins in the CREB/ATF family prefer a ten base-pair CRE target site (ATGACGTCAT) in which the same inverted pair of half sites is separated by two base pairs. The CRE and AP-1 target sites differ by just a single G•C base pair, yet they are the nuclear end-points of two different signal transduction pathways.[7]

We reported that the CRE target site recognized by CREB/ATF bZIP proteins bends intrinsically toward the major groove by 10-15°.[8] When CRE-BP1, one member of the CREB/ATF family, binds the CRE target site, it removes the intrinsic bend and effectively "straightens" the DNA; binding to the AP-1 site is accompanied by an induced bend toward the minor groove.[8] Based on these observations, we proposed two limiting models to account for the differential affinities of CREB/ATF family members for the CRE and AP-1 target sites.[9] In one

model, specificity is controlled by differential DNA flexibility;[10] the greater stability of protein•CRE complexes reflects a high free energy cost for bending the AP-1 target site. In a second model, specificity is controlled by differential protein•DNA contacts; the greater stability of protein•CRE complexes reflects a poor fit between the CREB/ATF protein and the AP-1 target site. Here we have studied the cyclic–AMP response-element binding protein (CREB),[11] another member of the CREB/ATF family, to assess the relative contributions of these two models. Of all well-characterized CREB/ATF family members, CREB is among the most divergent from CRE-BP1 in its primary sequence. CREB differs from CRE-BP1 at 14 of 26 positions within the basic and spacer segments, the region whose residues contact DNA directly (Figure 1). Here we show that the effect of CREB on the conformations of the CRE and AP-1 target sites is comparable to the effect of CRE-BP1. However, the relative affinities of CREB and CRE-BP1 for these two DNA sequences are markedly different. Taken together, these results indicate that CRE/AP-1 specificity, the hallmark of the CREB/ATF transcription factor family, cannot be explained solely on the basis of differential DNA flexibility, and must involve differential protein•DNA contacts. In addition, our data support a model in which DNA bending is controlled by a set of two basic residues located at the very NH_2-terminus of the basic segment (Figure 1).

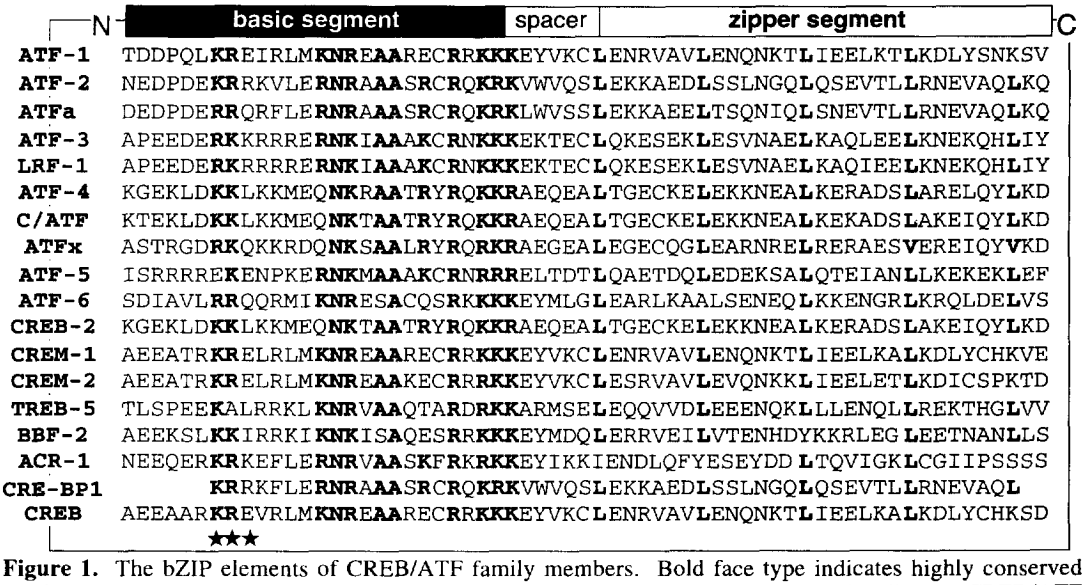

Figure 1. The bZIP elements of CREB/ATF family members. Bold face type indicates highly conserved residues. The residues above the stars have been proposed to be required for DNA bending by CREB/ATF family members.[8] Sequences were obtained from the following sources: ATF 1-6, ref. 12; ATFa, ref. 13; LRF-1, ref. 14; C/ATF, ref. 15; ATFx, ref. 16; CREB-2, ref. 17; CREM-1, ref. 18,19; CREM-2, ref. 19,20; TREB-5, ref. 21; BBF-2, ref. 22,23; ACR-1, ref. 24; CRE-BP1, ref. 25; CREB, ref. 11,26.

The CREB bZIP element displays high CRE/AP-1 specificity. The bZIP element of human CREB (residues 273-341), referred to here as bbb, was expressed from $pCREB_{259-327}$[27] in *E. coli* strain BL21(DE3)pLys S[28] and purified to homogeneity.[29] We used an electrophoretic mobility shift assay to quantify the affinity of bbb for the CRE and AP-1 target sites.[6,30] A typical binding experiment is illustrated in Figure 2A,

and binding isotherms showing the fraction of DNA bound (Θ) as a function of total bbb concentration are illustrated in Figure 2B.[31] In each case, the data fit an equation describing the formation of a 2:1 peptide•DNA complex.[6,30] The bbb•CRE complex exhibited an equilibrium dissociation constant (K_{app}) of (2.8 ± 0.3) x 10^{-18} M^2 corresponding to ΔG^o_{obs}= -23.9 kcal•mol^{-1} whereas the bbb•AP-1 complex exhibited K_{app} of (3.5 ± 0.4) x 10^{-17} M^2 (ΔG^o_{obs}= -22.4 kcal•mol^{-1}). Thus, under these conditions, bbb preferred the CRE target site to the AP-1 target site with a differential binding energy ($\Delta\Delta G^o_{obs}$) of 1.5 kcal•mol^{-1}. The CREB bZIP element bbb was less CRE-selective than was the CRE-BP1 bZIP element ccc, which displayed a differential binding energy ($\Delta\Delta G^o_{obs}$) of >2.3 kcal•mol^{-1} under identical conditions.[6,30]

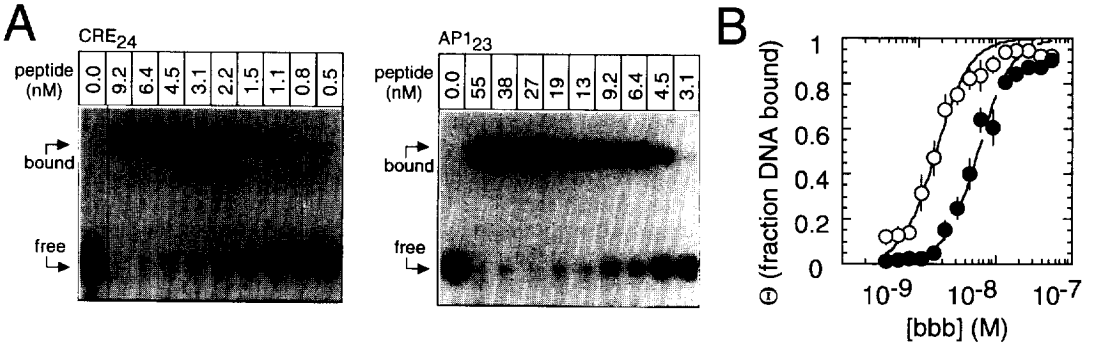

Figure 2. Equilibrium binding of CREB bZIP peptide (bbb) homodimers to DNA. **(A)** Autoradiogram illustrating the binding of ^{32}P end-labeled CRE_{24} and $AP\text{-}1_{23}$ to increasing concentrations of bbb at 25 °C. The total bbb concentration is shown above each lane. **(B)** Plots illustrating the fraction of DNA bound (Θ) as a function of $[bbb]_{total}$ concentration at 25 °C. The fraction CRE_{24} bound is depicted with open symbols; the fraction $AP\text{-}1_{23}$ bound is displayed with closed symbols. The curves represent the best fit of the data to the equation $\Theta = 1/(1+ K_{app}/[bbb]_{total}^2)$ where K_{app} is an adjustable parameter. The binding isotherms shown represent the average of at least 9 independent determinations. Error bars denote the standard error.[32]

$CREB_{273\text{-}341}$ straightens the intrinsically bent CRE target site. We employed a helical phasing analysis to determine the conformations of the CRE and AP-1 target sites when free and when in complex with bbb.[33] We used two sets of DNA test fragments containing a reference A-tract of defined curvature[34] separated from a CRE or AP-1 target site by a linker that varied the distance between the two sites in five steps over one helical turn.[35] As expected, the free CRE test fragments exhibited phase-dependent variations in mobility whereas the AP-1 test fragments did not [8] (Figure 3A). However, once bound to bbb, only the AP-1 test fragments exhibited phase-dependent changes in mobility – the CRE test fragments did not. The change in the relative mobility of the two sets of test fragments upon binding to bbb is clear evidence of a peptide-induced DNA bend. To determine the orientations of the induced bends, we plotted the relative mobility of each complex, uncorrected for the mobility of the free DNA, as a function of the distance in base pairs between the center of the A-tract and the test target site (Figure 3B). These plots show that, in both cases, the binding of bbb caused the DNA to bend toward the minor groove. In the case of the CRE target site, this minor groove bend neutralized the intrinsic major groove bend and effectively "straightened" the DNA (estimated bend angle[36] = 0°). In the case of the AP-1 site, the bend was toward the minor groove (estimated bend angle[36] = –6°). These results are consistent with previous studies on the conformation of the AP-1 target site in complex with a related CREB bZIP peptide.[36]

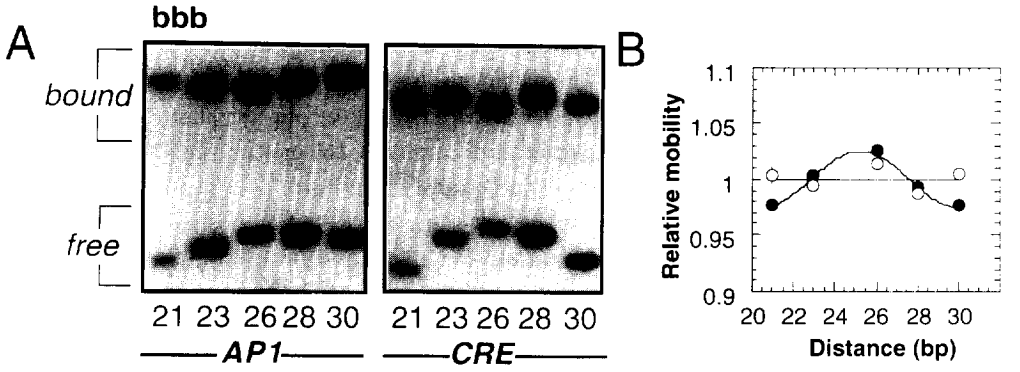

Figure 3. Phasing analysis of the conformations of the CRE and AP-1 target sites when in complex with the CREB bZIP peptide bbb. **(A)** Electrophoretic mobility-shift analysis[37,38] of bbb homodimers bound to phasing analysis probes.[8,33] **(B)** Relative mobilities plotted as a function of the distance in base pairs between the center of the test DNA site (CRE or AP-1) and the A-tract site. The CRE•bbb complexes are depicted with open symbols; the AP-1•bbb complexes are displayed with closed symbols. The mobilities of the bound DNA fragments were taken as the distance in millimeters from the center of the electrophoresis well to the center of the band corresponding to the bound DNA. These values were normalized to the average mobility of the fastest and slowest fragments to give relative mobilities. Complex mobilities were not corrected for the changes in the mobilities of the free DNA test fragments because of the intrinsic bend in the CRE target site. The relative mobilities of the complexes represent the average of at least 3 independent experiments. Error bars indicating the standard error are shown.[32] The points are connected by the calculated best fit of the data to a cosine curve.[36]

Implications for the structural determinants of CRE straightening. Several other bZIP family members bend the AP-1 target site when they bind.[36] We compared the sequences of these bZIP proteins and noticed a correlation between certain residues in the basic segment and the direction of protein-induced bending.[8] Proteins that induce a minor groove bend contain three adjacent basic residues (Lys-Arg-Arg or Arg-Arg-Arg) at the amino terminus of the basic segment, whereas proteins that induce a major groove bend contain a conserved pattern of charged and hydrophobic residues in these positions. GCN4, which does not induce a DNA bend, contains the sequence Pro-Ala-Ala and this sequence does not contact DNA.[2,3] [The locations of these residues are indicated with stars in Figure 1 underneath the sequence of CRE-BP1.] On the basis of these comparisons, we proposed that the differential interactions of these basic segment residues with DNA may regulate the differential bending and binding of bZIP proteins.[8] CREB contains only two of these residues, yet it is still able to straighten the CRE target site (Figure 1). This result suggests that only two basic residues (four per complex) may be necessary to distort the CRE target site into the a conformation that is more appropriate for binding. It is interesting to note that these two basic residues are entirely conserved within members of the CREB/ATF family, consistent with the idea that the function of CREB/ATF proteins, that is, the preferential recognition of CRE target sites in response to cAMP, requires a protein-induced DNA bend.

Implications for the origin of CRE/AP-1 specificity. Previously we proposed two limiting models to account for the CRE/AP-1 target-site specificity of CREB/ATF family members (Figure 4).[8] In Model 1, the stability of CRE•bZIP complexes when compared with AP-1•bZIP complexes reflects a high cost for bending the

AP-1 target site. Here, the energy required to distort the intrinsically straight AP-1 target site is greater than that required to straighten the intrinsically bent CRE target site; the energy gained upon formation of the two protein•DNA complexes is the same. Specificity in this case is determined solely by the differential energy cost of bending the CRE and AP-1 target sites; this model therefore predicts that all CREB/ATF proteins will exhibit identical CRE/AP-1 specificities. In Model 2, the higher stability of CRE•protein complexes reflects a poor fit between the recognition interface of the CREB/ATF protein and the AP-1 target site. Here, the energy required to distort the intrinsically straight AP-1 target site is equal to the energy required to straighten the intrinsically bent CRE target site; the free energy gained upon formation of the CRE•bZIP complex is greater than that gained upon formation of the AP-1•bZIP complex. This model predicts that the CRE/AP-1 specificities of CREB/ATF proteins may vary.

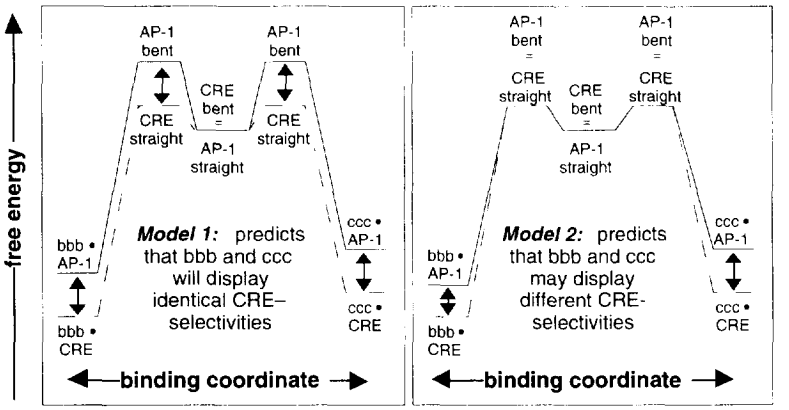

Figure 4. Two limiting models to account for the preference of CREB/ATF proteins for the CRE target site. Shown are the relative free energies of the AP-1 and CRE target site and their complexes with the CREB bZIP peptide bbb and the CRE-BP1 bZIP peptide ccc. The energies of the intrinsically bent CRE target site and the straight AP-1 target site are considered to be equal in both models.

Here we show that the effect of the CREB bZIP peptide on the conformations of the CRE and AP-1 target sites is comparable to that of the CRE-BP1 bZIP peptide. Both peptides remove the intrinsic major groove bend in the CRE target site and induce a comparable minor groove bend in the AP-1 target site. However, the CRE-BP1 bZIP ccc is more selective for the CRE target site than is the CREB bZIP bbb by a differential $\Delta\Delta G^{o}_{obs}$ of almost 1 kcal•mol^{-1}.[6] The observation that CREB induces the same DNA distortion as CRE-BP1 but displays markedly lower CRE/AP-1 specificity places a 1.5 kcal•mol^{-1} upper limit on the differential binding energy that may be derived from differential DNA flexibility. Further work will be necessary to define precisely how much models 1 and 2 contribute to the control of specificity in these proteins, and how accessory factors that bind these protein complexes *in vivo* modulate specificity.

Acknowledgment. This work was supported by the NIH. A.S. is an Eli Lilly Fellow, a David and Lucile Packard Foundation Fellow, a Camile and Henry Dreyfus Teacher–Scholar, an NSF Presidential Young Investigator, an Alfred P. Sloan Research Fellow, and a 1995 Arthur C. Cope Scholar. We are grateful to Dr. Ourania Andrisani for the generous gift of pCREB$_{279\text{-}341}$.

References and Notes

1. McKnight, S. L. *Sci. Am.* **1991**, *264*, 54.
2. Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223.
3. König, P.; Richmond, T. *J. Mol. Biol.* **1993**, *233*, 139.

4. Harrison, S. C. *Nature (London)* **1991**, *353*, 715.
5. Glover, J. N. M.; Harrison, S. C. *Nature (London)* **1995**, *373*, 257.
6. Metallo, S. J.; Schepartz, A. *Chemistry & Biology* **1994**, *1*, 143.
7. Habener, J. F. *Mol. Endocrin.* **1990**, *4*, 1087.
8. Paolella, D. N.; Palmer, C. R.; Schepartz, A. *Science* **1994**, *264*, 1130.
9. Schepartz, A. *Science* **1995**, in press.
10. Klug, A.; Jack, A.; Viswamitra, M. A.; Kennard, O.; Steitz, T. A. *J. Mol. Biol.* **1979**, *131*, 669.
11. Hoeffler, J. P.; Meyer, T. E.; Yun, Y.; Jameson, J. L.; Habener, J. F. *Science* **1988**, *242*, 1430.
12. Hai, T.; Liu, F.; Coukos, J.; Green, M. R. *Genes Dev.* **1989**, *3*, 2083.
13. Gaire, M.; Chatton, B.; Kedinger, C. *Nuc. Acids Res.* **1990**, *18*, 3467.
14. Hsu, J. C.; Laz, T.; Mohn, K. L.; Taub, R. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 3511.
15. Vallejo, M.; Ron, D.; Miller, C. P.; Habener, J. F. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4679.
16. Nishizawa, M.; Nagata, S. *FEBS Lett.* **1992**, *299*, 36.
17. Karpinski, B. A.; Morle, G. D.; Huggenvik, J.; Uhler, M. D.; Leiden, J. M. *Proc. Natl. Acad. Sci USA* **1992**, *89*, 4820.
18. Masquilier, D.; Foulkes, N. S.; Mattei, M. G.; Sassoni-Corsi, P. *Cell Growth Differ.* **1993**, *4*, 931.
19. Foulkes, N. S.; Borrelli, E.; Sassone-Corsi, P. *Cell* **1991**, *64*, 739.
20. Meyer, T. E.; Habener, J. F. *Nuc. Acids Res.* **1992**, *20*, 6106.
21. Yoshimura, T.; Fujisawa, J.-i.; Yoshida, M. *EMBO J.* **1990**, *62*, 2537.
22. Abel, T.; Bhatt, R.; Maniatis, T. *Genes Dev.* **1992**, *6*, 466.
23. Smolik, S. M.; Rose, R. E.; Goodman, R. H. *Mol. Cell. Biol.* **1992**, *12*, 4123.
24. Vincent, A. C.; Struhl, K. *Mol. Cell. Biol.* **1992**, *12*, 5394.
25. Maekawa, T.; Sakura, H.; Kanei-Ishii, C.; Sudo, T.; Yoshimura, T.; Fujisawa, J.; Yoshida, M.; Ishii, S. *EMBO J.* **1989**, *8*, 2023.
26. Gonzalez, G. A.; Yamamoto, K. K.; Fischer, W. H.; Karr, D.; Menzel, P.; III, W. B.; Vale, W. W.; Montimony, M. R. *Nature (London)* **1989**, *337*, 749.
27. Santiago-Rivera, Z. I.; Williams, J. S.; Gorenstein, D. G.; Andrisani, O. M. *Protein Sci.* **1993**, *2*, 1461.
28. Studier, F. W.; Moffat, B. A. *J. Mol. Biol.* **1985**, *189*, 113.
29. Cells were grown to an optical density of 0.5 at 600 nm, induced with 0.5 mM IPTG for 4 hours, harvested, resuspended in 10 mM sodium phosphate (pH 7.4), 0.5 M NaCl, 1 mM EDTA, and 5 mM DTT, and lysed by sonication. The supernatant was heated at 75 °C for 10 min, centrifuged, and the peptide was precipitated from the supernatant with 80% $(NH_4)_2SO_4$ at 4 °C. The precipitate was dialyzed into 10 mM phosphate (pH 7.4) and 5 mM DTT. After dialysis, the peptide was purified to homogeneity on hydroxyapatite HT Bio-Gel (BioRad) resin equilibrated with 10 mM phosphate (pH 7.4), 0.2 M NaCl, and 5 mM DTT and was washed with 50 mM phosphate (pH 7.4), 0.5 M NaCl, and 5 mM DTT. The peptide was stored in 1.3 mM phosphate (pH 7.4), 13 mM NaCl, and 5 mM DTT. Final peptide concentrations were determined spectroscopically and by amino acid analysis and were corrected for activity (20%).
30. Metallo, S. J.; Schepartz, A. In *Tech. Protein Chem.*; J. W. Crabb, Ed.; Academic Press: San Diego, 1995; Vol. VI.
31. Binding reactions were performed in phosphate-buffered saline (PBS) buffer [137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4 (pH 7.4)] containing 1 mM EDTA, 1 mM DTT, 0.1% NP40, 0.4 mg/mL BSA, and 5% glycerol. Peptide was added to ^{32}P end-labeled DNA (<50 pM) and the binding reactions were incubated for 30 min at 25 °C and then loaded onto a 16 x 18 cm non-denaturing 8% (80:1 acrylamide: bisacrylamide) gel prepared in 45 mM Tris-borate, 1 mM EDTA, pH 8.0. The electrophoresis running buffer contained 22.5 mM Tris-borate, 0.5 mM EDTA, pH 8.0. Gels were pre-equilibrated for 40 min at 500 V and were maintained at 25 °C by immersion in a circulating, temperature-controlled water bath. The gels were run for 300 V•H. The amount of complexed and free DNA was quantified on a Betagen 603 Blot Analyzer. $\Delta G°obs$ values were calculated from the equation $\Delta G°obs = -RTln(1/Kapp)$ where R=0.001987 kcal•mol^{-1}•K^{-1} and T=298K. CRE_{24}: d(AGTGGAG**ATGACGTCAT**CTCGTGC); $AP\text{-}1_{23}$: d(AGTGGAG**ATGACTCATC** TCGTGC).
32. Devore, J. L. *Probability and Statistics for Engineering and the Sciences*; Brooks/Cole: Monterey, 1987.
33. Zinkel, S. S.; Crothers, D. M. *Nature (London)* **1987**, *328*, 178.
34. Koo, H.-S.; Drak, J.; Rice, A. J.; Crothers, D. M. *Biochemistry* **1990**, *29*, 4227.
35. Binding reactions were performed by incubating 20 to 30 nM of peptide with 10 to 100 pM of ^{32}P end-labeled double-stranded DNA fragment for 30 min at 4 °C in PBS buffer containing 1 mM EDTA, 1 mM DTT, 0.05% NP40, and 5% glycerol. The reactions were loaded onto a non-denaturing 8% (32:1) polyacrylamide gel in TG buffer (18.75 mM Tris (pH 8.9), 146.25 mM glycine). Gels were run for 1600 V•H and were maintained at 4 °C as described above.
36. Kerppola, T. K.; Curran, T. *Mol. Cell. Biol.* **1993**, *13*, 5479.
37. Fried, M.; Crothers, D. M. *Nuc. Acids Res.* **1981**, *9*, 6505.
38. Garner, M. M.; Revzin, A. *Nuc. Acids Res.* **1981**, *9*, 3047.

(*Received in USA* 20 *April* 1995; *accepted* 13 *June* 1995)