17. E. J. Milner-White, *J. Mol. Biol.* **199**, 503 (1988).
18. E. N. Baker and R. E. Hubbard, *Prog. Biophys. Mol. Biol.* **44**, 97 (1984); R. Preißner and P. Bork, *Biochem. Biophys. Res. Commun.* **180**, 660 (1991); P. Bork and R. Preißner, *ibid.*, p. 666; S. Dasgupta and J. A. Bell, *Int. J. Peptide Protein Res.* **41**, 499 (1993); H. A. Nagarajram, R. Sowdhamini, C. Ramakrishnan, P. Balaram, *FEBS Lett.* **321**, 79 (1993).
19. J. S. Richardson, *Adv. Protein Chem.* **34**, 167 (1981).
20. In the Schellman motif, the mean accessibility of N and O is 7% in the (C″) N–H. . . O=C (C3) hydrogen bond and 25% in the (C′) N–H . . . O=C (C2) hydrogen bond. Correspondingly, in the α_L motif, the mean accessibility of N and O is 9% in the (C′) N–H . . . O=C (C3) hydrogen bond. Solvent-accessible surface area is defined by B. K. Lee and F. M. Richards [*J. Mol. Biol.* **55**, 379 (1971)] and calculated with ANAREA, an algorithm attributable to T. J. Richmond [*ibid.* **178**, 63 (1984)]. Results were normalized to a percentage with the use of values from G. J. Lesser and G. D. Rose [*Proteins Struct. Funct. Genet.* **8**, 6 (1990)].
21. K. R. Shoemaker *et al.*, *Nature* **326**, 563 (1987).
22. M. Schiffer and A. B. Edmundson, *Biophys. J.* **7**, 121 (1967).
23. M. J. Cocco and J. T. J. Lecomte, *Protein Sci.*, in press.
24. Backbone dihedral angles used for C3–C2–C1 were $\phi = -64°$, $\psi = -41°$; van der Waals radii were taken from (*34*).
25. All conformations of the hexapeptide Leu-Ala-Ala-Ala-Gly-Ile were sampled on a 30° grid, with the first three residues held helical. Side chain torsions were allowed to occupy only staggered conformations. The resulting 291,599 allowed conformations constituted the "working set." To assess the role of hydrogen bonds (*35*) in a Schellman motif, the working set was filtered to select conformations in which hydrogen-bonded partners were within 3.5 Å. Upon energy minimization of the 157 conformations satisfying this criterion, all conformations with hydrogen bonds also had the two hydrophobic residues in contact. Minimizations were performed with the AMBER/OPLS force field with MacroModel (*34*). In the converse experiment, all conformations with at least one hydrophobic contact between residues 1 and 6 were selected and energy-minimized. The most stable structure was the Schellman motif. Because the Schellman and α helix structures contain the same number of hydrogen bonds, the factor stabilizing the Schellman over the helix is probably the hydrophobic contact. Consistent with this view, energy minimization of the model peptide that was started from a helical conformation resulted in considerable distortion, but the initial structure persisted when started from a Schellman conformation. The number of hydrogen bonds and hydrophobic contacts is identical for an α_L motif and an α helix of the same length, so there is no compelling reason for helix termination over continuation. A possible explanation is the greater conformational freedom available to Gly in the less restrictive α_L conformation. Other residues can substitute for Gly in the Schellman motif but not in the α_L motif, perhaps because the strain imposed by the adoption of a right-handed conformation is counterbalanced by the existence of favorable hydrogen bond and hydrophobic contacts between residues C3 and C″ in a Schellman motif. Accordingly, energy minimization showed that a Gly-based α_L is less disfavored than its Ala-based counterpart.
26. U. Hobhorn, M. Scharf, R. Schneider, C. Sander, *Protein Sci.* 1, 409 (1992), as applied to the November 1993 release of the PDB (*13*).
27. T. P. Creamer and G. D. Rose, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5937 (1992); S. D. Picket and M. J. E. Sternberg, *J. Mol. Biol.* **231**, 825 (1993).
28. P. S. Kim and R. E. Baldwin, *Nature* **307**, 329 (1984).
29. P. J. Loll and E. E. Lattman, *Proteins Struct. Funct. Genet.* **5**, 183 (1989).
30. D. Shortle *et al.*, *Biochemistry* **29**, 8033 (1990); S. M. Green *et al.*, *ibid.* **31**, 5717 (1992).
31. S. J. Remington *et al.*, *J. Mol. Biol.* **118**, 81 (1978).
32. T. Alber *et al.*, *Nature* **330**, 41 (1987).
33. A. Horowitz *et al.*, *J. Mol. Biol.* **219**, 5 (1991); L. Serrano *et al.*, *ibid.* **224**, 847 (1992).
34. W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110**, 1657 (1988); S. J. Weiner *et al.*, *ibid.* **106**, 765 (1984); F. Mohamadi *et al.*, *J. Comput. Chem.* **4**, 440 (1990).
35. D. F. Stickle, L. G. Presta, K. A. Dill, G. D. Rose, *J. Mol. Biol.* **226**, 1143 (1992).
36. Computer-generated drawings were made with Insight II, 2.30; Biosym Technologies, San Diego, CA. Illustrations were provided with SCIAN, 0.852B; E. Pepke, J. Murray, J. Lyons, and T.-Y. Hwo, Supercomputer Computations Research Institute, Florida State University, Tallahassee, FL.
37. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
38. We thank T. Creamer, E. Harper, N. Kallenbach, E. Lattman, and J. Seale for useful discussions; J. Ponder for the area subroutine from his TINKER package; R. Baldwin, N. Kallenbach, and J. Lecomte for preprints of papers; C. Frieden for a critical reading of the manuscript; and K. Henrick and A. Fersht for barnase coordinates. Supported by a grant from the NIH (GM 29458).

7 December 1993; accepted 22 March 1994

DNA Targets for Certain bZIP Proteins Distinguished by an Intrinsic Bend

David N. Paolella, C. Rodgers Palmer, Alanna Schepartz*

In spite of the large amount of sequence conservation among the DNA binding segments of basic region leucine zipper (bZIP) proteins, these proteins can discriminate differently between target sequences that differ in half-site spacing. Here it is shown that the half-site spacing preferences of bZIP proteins are the result of (i) the differential intrinsic curvature in target binding sites that differ by insertion or deletion of a single base pair and (ii) the ability of some bZIP proteins to overcome this intrinsic curvature through a mechanism dependent on basic segment residues.

The bZIP family of eukaryotic transcription factors contains (i) a short, helical basic segment whose residues participate in DNA contacts, (ii) a zipper segment responsible for protein dimerization (*1*, *2*), and (iii) a six-residue spacer of variable sequence connecting the two (*3*). X-ray crystallography data from two different bZIP-DNA complexes show that the dimeric bZIP protein contains a pair of uninterrupted α helices that interact with each other along the length of the zipper segment to form a parallel coiled coil (*4*). These α helices diverge in the vicinity of the nucleic acid and interact with the major groove of the target DNA (*5*, *6*).

Like certain other dimeric DNA binding proteins (*7*, *8*), bZIP proteins discriminate differently between target binding sequences that differ only in half-site spacing. For example, proteins related to Fos and Jun (AP-1 family) bind preferentially to a pseudosymmetric 9–base pair (bp) site, which consists of two ATGA half-sites arranged in an inverted pair separated by a single dC:dG base pair—that is, ATGACTCAT. Proteins related to cyclic adenosine monophosphate response element–binding protein (CREB) and to activating transcription factor–2 (ATF-2) (CREB-ATF family) have higher affinity for the dyad-symmetric CRE site in which the same inverted pair of half-sites is separated by 2 bp (ATGACGTCAT) (*9*). In contrast to the AP-1 and CREB-ATF families, the yeast bZIP protein GCN4 binds to both sites with comparable affinity (*10*). Within the context of B DNA, the additional dG:dC base pair in the CRE site displaces the two ATGA contact surfaces by an axial translation of 3.4 Å and a twist angle of 34° (*5*). The ability of GCN4 to accept both sequences as specific targets could be the result of the flexibility of an α-helical segment, which permits a structural readjustment that its counterpart in CREB and ATF proteins does not permit (*6*, *11*). Alternatively, GCN4 could bind with the same structure but there could be an induced structural readjustment of one or both of the DNA targets (*5*). A third possibility is that GCN4 has an equal affinity for the two DNA target sites as a result of an intrinsic deformation of the B-form structure of one target site which compensates for the difference in half-site spacing.

Initially, we used a circular permutation assay (*12*) to compare the conformation of the CRE and AP-1 target DNA sequences both alone and when bound to the bZIP segments of GCN4 and CRE-BP1, a member of the CREB-ATF family; we have labeled these bZIP peptides ggg and ccc, respectively (Fig. 1). The circular permutation assay is based on the position-dependent effects of DNA distortion on the electrophoretic mobilities of a set of isomeric DNA fragments (*12*). Oligonucleotides with conformational distortions near

D. N. Paolella and A. Schepartz, Department of Chemistry, Training Program in Biophysics, Yale University, New Haven, CT 06511, USA.
C. R. Palmer, Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA.

*To whom correspondence should be addressed.

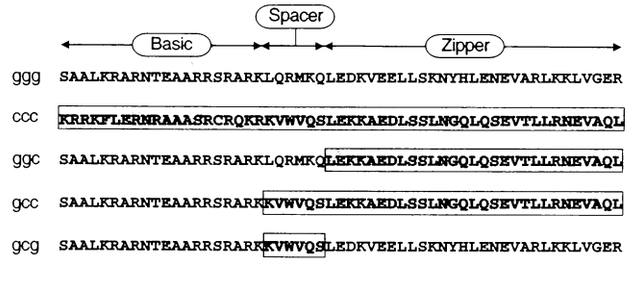

Fig. 1. Amino acid sequences of the bZIP peptides and chimeras used here (*27*). Peptide ggg contains an NH_2-terminal serine followed by residues 228 to 281 of GCN4 (*28*); peptide ccc contains residues 354 to 408 of CRE-BP1 (*29*). Shading indicates amino acid sequence derived from CRE-BP1. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

the center of the DNA fragment show anomalously slow mobility in a nondenaturing polyacrylamide gel when compared with oligonucleotides of similar length with distortions near the end of the DNA fragment. All of the bZIP-DNA complexes tested in this assay displayed position-dependent variations in electrophoretic mobility (Fig. 2), whereas the unbound DNA fragments possessed uniform mobility regardless of the position of the target sequence within the fragment (*13*).

We investigated the apparent DNA distortion further using an analysis based on the phase-dependent effects on electrophoretic mobility, which are caused when a DNA fragment containing the target sequence also contains a reference sequence of defined curvature (*14, 15*). Phasing analysis is more sensitive than circular permutation analysis, and it distinguishes distortions as a result of directed bends from those that result from isotropic flexibility or other distortions. It also defines the orientation of a directed bend relative to the reference standard incorporated into the fragment. Two sets of oligonucleotides were constructed that contained the CRE or AP-1 target sequence separated by a variable length linker from a 25-bp sequence that contained an A tract of defined curvature (*16*) (Fig. 3A). As expected (*5, 6, 17*), the ggg-CRE complexes (but not the ggg–AP-1 complexes) showed phase-dependent variations in electrophoretic mobility (Fig. 3B). To determine the orientation of the bend, we plotted the relative mobilities of the ggg-CRE complexes as a function of the distance in base pairs between the centers of the two sites (Fig. 3C). The construct with the slowest mobility contained a center-to-center spacing of 26 bp, or two and a half helical turns (Fig. 3C). This result indicates that in solution, like in the crystal (*5*), the center of the CRE binding sequence is bent toward the major groove in the complex—that is, bent toward the zipper. Analysis (*15, 18, 19*) of the differences in mobility exhibited by the ggg-CRE complexes suggests that there is a bend angle of 13°, an angle close to the 20° bend observed in the x-ray structure (*5*).

Even in the absence of the ggg-bZIP peptide, the CRE-containing oligonucleotides exhibited variations in mobility that were phase-dependent. Among the free CRE constructs, the one with the slowest mobility contained a CRE target site with a center two and a half helical turns from the center of the A tract (Fig. 3C). No phase-dependent variations in mobility were observed with the AP-1 constructs, which differed from the CRE constructs by deletion of 1 bp. To evaluate the extent of CRE curvature in the presence and absence of ggg, we compared the phase-dependent variations in mobility in the presence and absence of bound peptide. We define R as the ratio of the normalized mobilities for the ggg-CRE complexes divided by the corresponding free CRE mobilities [$\mu_{b(rel)}/\mu_{f(rel)}$]. The value of R was within 2% of unity and independent of the phasing between the CRE and A tract sequences (Fig. 3D). This result indicated that the change in mobility of the ggg-CRE complexes that occurred as the spacing between the center of the CRE binding site and the center of the A tract was changed was not caused by binding of the ggg homodimer. The lack of change in the conformation of the CRE site upon binding of ggg implies that the CRE sequence is bent intrinsically into a conformation suitable for ggg. Analysis of the variations in mobility exhibited by the CRE site estimated an intrinsic bend of 11° (*15, 18–20*), compared with a bend angle of 3° that was estimated for the free AP-1 sequence.

The crystal structure of the GCN4-CRE complex shows that DNA bending and unwinding places the two dyad-symmetric ATGA half-sites in an orientation with respect to each other that is approximately the same as that observed in the GCN4–AP-1 complex (*5*). Our observations indicate that these structural adjustments in the CRE site are not imposed by the bound GCN4 homodimer but are encoded in the sequence. Because no binding energy is lost in the deformation of the CRE site, our observations explain why GCN4 binds equally well to the undeformed AP-1 site and the bent CRE site. No major changes in the structure of GCN4 (*6, 11*), in the CRE or AP-1 half-sites (*5*), or in the interface between the GCN4 basic region and the target half-site are required to explain the target site preferences of GCN4.

The observation that the bZIP binding surface of the 10-bp CRE sequence is deformed intrinsically and thus resembles that of the 9-bp AP-1 target sequence poses a paradox. If the native conformations of the CRE and AP-1 target sequences present similar constellations of functional groups in the major groove so that GCN4 can recognize both, it would be difficult to understand how CREB and ATF proteins discriminate between them. Phasing analysis revealed that although the CRE sequence displayed major groove curvature in the absence of ccc, it showed no curvature when bound to ccc (Fig. 3, B and C). The bend angle estimated for the ccc-CRE complex is 4°, an angle similar to that estimated for the free AP-1 sequence (3°). The removal of intrinsic major groove curvature in the CRE site upon binding of ccc requires a protein-induced bend toward the minor groove and probably other structural adjustments that straighten the DNA. Although ccc bound the CRE sequence with approximately 50 times greater affinity than it bound the AP-1 sequence with (*21*), a ccc–AP-1 complex formed at high protein concentration showed clear evidence of bending toward the minor groove (Fig. 3, B and C).

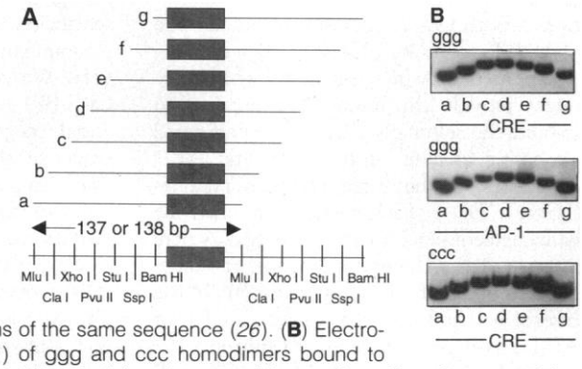

Fig. 2. Circular permutation analysis (*12*) of distortion in the CRE and AP-1 complexes of ggg and ccc homodimers. (**A**) The probes used for circular permutation analysis were generated by restriction endonuclease cleavage at sites within two tandem polylinker sequences flanking the CRE (or AP-1) sites (shaded box), with the seven enzymes shown (a through g) (*30*). All probes were 138 bp (137 bp for AP-1) in length and contained circular permutations of the same sequence (*26*). (**B**) Electrophoretic mobility-shift analysis (*31*) of ggg and ccc homodimers bound to circularly permuted probes (*32*). Peptide ccc did not form a stable complex with probes containing the AP-1 site.

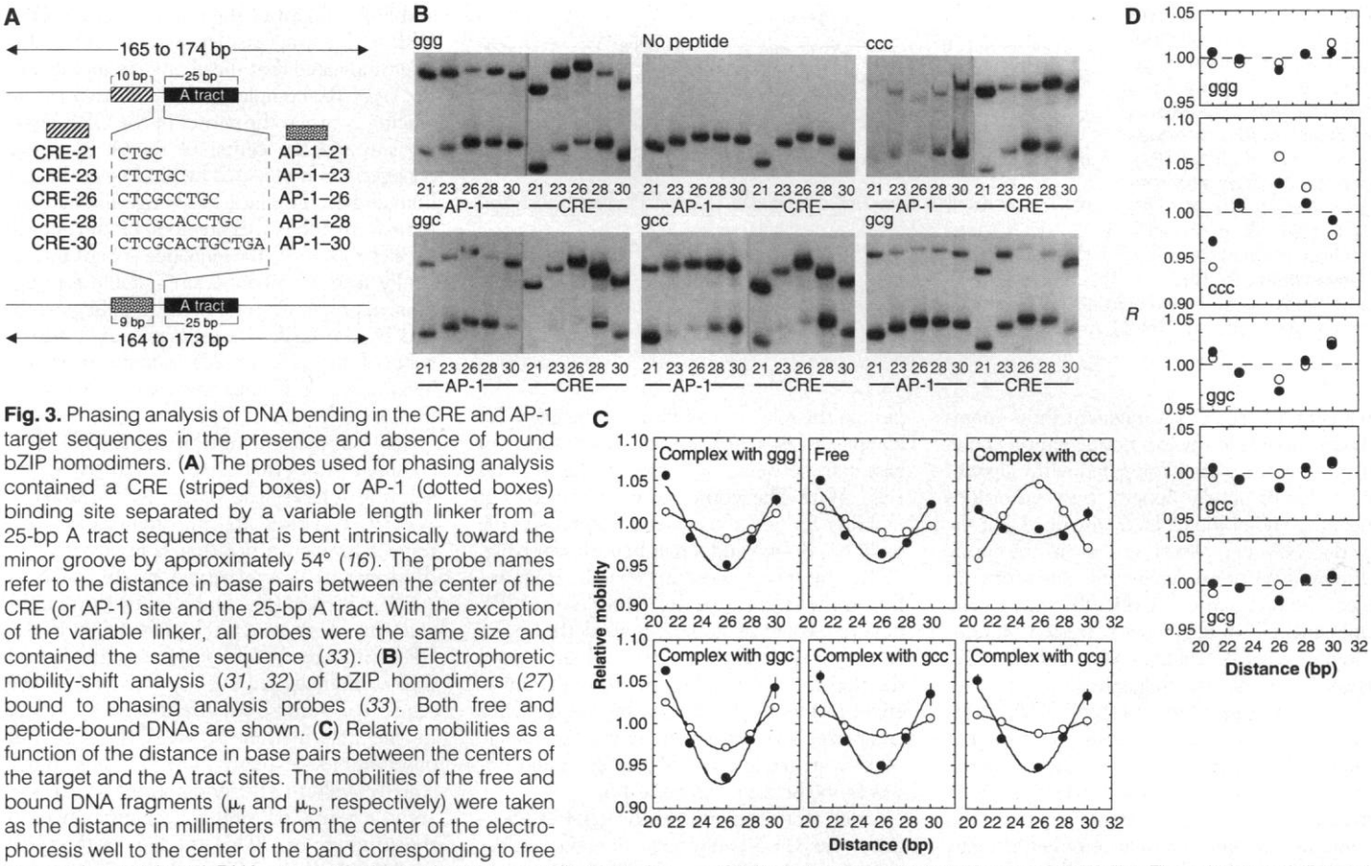

Fig. 3. Phasing analysis of DNA bending in the CRE and AP-1 target sequences in the presence and absence of bound bZIP homodimers. (**A**) The probes used for phasing analysis contained a CRE (striped boxes) or AP-1 (dotted boxes) binding site separated by a variable length linker from a 25-bp A tract sequence that is bent intrinsically toward the minor groove by approximately 54° (*16*). The probe names refer to the distance in base pairs between the center of the CRE (or AP-1) site and the 25-bp A tract. With the exception of the variable linker, all probes were the same size and contained the same sequence (*33*). (**B**) Electrophoretic mobility-shift analysis (*31, 32*) of bZIP homodimers (*27*) bound to phasing analysis probes (*33*). Both free and peptide-bound DNAs are shown. (**C**) Relative mobilities as a function of the distance in base pairs between the centers of the target and the A tract sites. The mobilities of the free and bound DNA fragments (μ_f and μ_b, respectively) were taken as the distance in millimeters from the center of the electrophoresis well to the center of the band corresponding to free or protein-complexed DNA, respectively, and were normalized to the average mobility of the fastest and slowest fragments to give relative mobilities $\mu_{f,n}$ and $\mu_{b,n}$, respectively ($\mu_{b,n} = \mu_b/\mu_{b(avg.)}$; $\mu_{f,n} = \mu_f/\mu_{f(avg.)}$). The complex mobilities were not corrected for changes in the mobilities of the free probes because of the intrinsic bend in the CRE sequence. The curvature in the CRE test fragments was not the result of intrinsic curvature of the variable linker, because both sets of CRE (●) and AP-1 (○) test duplexes contained identical linkers, and only the CRE set exhibited phase-dependent changes in mobility. The relative mobilities of the complexes represent the average of at least four independent experiments (except for the ccc–AP-1 complex). Error bars represent one standard deviation. The points are connected by the calculated best fit of the data to a cosine function (*19*). (**D**) Ratio of the normalized mobilities for the CRE (●) and AP-1 (○) complexes of bZIP homodimers; *R* is defined as the ratio $\mu_{b,n}/\mu_{f,n}$. Distance is defined as the distance in base pairs between the centers of the CRE (or AP-1) site and the A tract.

We interpret these findings: A bend toward the major groove and an accompanying unwinding cause the dyad-symmetric CRE target sequence to diminish its apparent half-site spacing. The intrinsic bend toward the major groove compensates structurally for the 2-bp spacing between half-sites and mimics the 1-bp spacing of the AP-1 site. For either DNA, a bend toward the minor groove and an accompanying overwinding cause the opposite effect: an increase in both the axial displacement of the half-sites and the helical twist angle. Thus, the return to unbent DNA imposed on the CRE site by ccc binding restores the axial and azimuthal separation of 2 bp that is appropriate for ccc. If the AP-1 site with a 1-bp spacing between half-sites is bent toward the minor groove, then it should mimic the target with a larger spacing between half-sites and should be a suitable target for ccc. The difference is that the recognition interface of ccc, the natural cognate of the CRE site, appears designed to straighten the inherently bent CRE site with little loss in binding energy, but it is not well designed to deform the AP-1 site. Either the energy required to induce a minor groove bend in the AP-1 sequence is higher than the energy required to remove the major groove bend in the CRE sequence, or stabilizing interactions between the (straight) CRE sequence and ccc are not in register in the complex with the (bent) AP-1 site; alternatively, a combination of both factors operates to produce the observed specificity (*22*) (Fig. 4).

To identify which segments within ccc might provide the interactions required to stabilize the deformed CRE site, we performed a phasing analysis of the CRE and AP-1 complexes of the three chimeric peptides shown in Fig. 1. Each peptide contained the basic segment sequence of GCN4 (gxx, where x is either g or c) and the spacer or zipper segment of either GCN4 or CRE-BP1. In the case of gcc and gcg, the extent of CRE or AP-1 curvature in the complex mirrored the extent of curvature in the free DNA (Fig. 3, B and D). In the case of ggc, which contains only the zipper segment of CRE-BP1, a modest increase in major groove curvature was observed. The absence of induced minor groove curvature in the CRE or AP-1 complexes of any chimeric peptide that lacked the CRE-BP1 basic segment suggests that residues within the CRE-BP1 basic segment stabilize the minor groove bend induced by ccc (*23*). These results are consistent with those indicating that the determinants of CRE–AP-1 specificity, as measured by affinity, are found predominantly within the ccc basic segment (*21*). We conclude that residues within the CRE-BP1 basic segment regulate the preferential recognition of the dyad-symmetric CRE sequence through interactions with DNA that compensate, in whole or in part, for losses in free energy sustained through DNA distortions.

Several bZIP proteins induce a bend in the AP-1 sequence when they bind (*15, 18, 19*). On the basis of the observation that hetero- and homodimeric complexes of bZIP proteins induce bends in opposite directions, it was proposed that each protein monomer induced

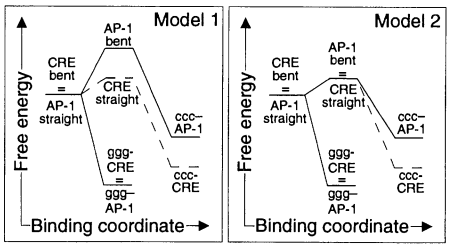

Fig. 4. Two limiting models to account for the preference of the ccc peptide for the CRE target site. Shown are the relative free energies of the AP-1 target sequence and its protein complexes (solid lines) and the CRE target sequence and its protein complexes (dotted lines) along the binding coordinate. The energies of the native CRE (bent) and AP-1 (straight) sequences are considered to be equal in both models. In model 1, the greater stability of the ccc-CRE complex reflects the difficulty of bending the AP-1 sequence. Here, the (bending) energy required to distort the intrinsically straight AP-1 sequence is greater than the (bending) energy required to straighten the intrinsically bent CRE sequence. The (binding) energy gained upon formation of the protein-DNA complex is the same in both cases. In model 2, the greater stability of the ccc-CRE complex reflects a poor fit between the recognition interface of ccc and the (bent) AP-1 sequence. Here, the (bending) energy required to distort the intrinsically straight AP-1 sequence is equal to the (bending) energy required to straighten the intrinsically bent CRE sequence. The (binding) energy gained upon formation of the ccc–AP-1 (bent) complex is less than that gained upon formation of the ccc-CRE (straight) complex.

an independent DNA bend (*15, 18, 19*). We localized the CRE-BP1 residues that induce bending to the basic segment. A comparison of bZIP proteins reveals a correlation between basic segment sequence and the direction of protein-induced bending. Proteins that induce a minor groove bend [CRE-BP1, ATF-2, Fos, and Fra (*15, 18, 19*)] contain three basic residues (Lys-Arg-Arg or Arg-Arg-Arg) at the very NH_2-terminus of the basic segment, whereas proteins that induce a major groove bend [Jun and Zta (*15, 18, 19*)] contain a conserved pattern of charged and hydrophobic residues in these positions. GCN4, which does not induce a DNA bend (*17*), contains the sequence Pro-Ala-Ala. Although these basic residues likely contact the DNA far from the site of curvature, it has been pointed out that interactions that stabilize a bend achieve maximum leverage when located far from the bend point (*24*). We propose that the differential interactions of these basic segment residues with DNA regulate the differential bending and binding of bZIP proteins.

Previous studies of DNA-protein interactions have established two mechanisms by which sequence specificity is achieved (*25*). One involves hydrogen bonding and van der Waals interaction of DNA base pairs with a matrix of protein side chains and water molecules that complements only the correct DNA sequence. The second involves the sequence-dependent ability of DNA to distort upon protein binding with minimal (or compensated) free energy cost. Here, we show that specificity in the formation of bZIP-DNA complexes is achieved in a third way, by pre-organization of the DNA into a distorted structure bound by only the appropriate bZIP protein.

REFERENCES AND NOTES

1. S. L. McKnight, *Sci. Am.* **264**, 54 (April 1991).
2. W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science* **240**, 1759 (1988).
3. W. T. Pu and K. Struhl, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6901 (1991).
4. E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *Science* **254**, 539 (1991).
5. P. König and T. Richmond, *J. Mol. Biol.* **230**, 139 (1993).
6. T. E. Ellenberger *et al.*, *Cell* **71**, 1223 (1992).
7. K. Umesono and R. M. Evans, *ibid.* **57**, 1139 (1989).
8. R. J. Reece and M. Ptashne, *Science* **261**, 909 (1993).
9. T. Hai *et al.*, *Genes Dev.* **3**, 2083 (1989).
10. J. W. Sellers, A. C. Vincent, K. Struhl, *Mol. Cell. Biol.* **10**, 5077 (1990).
11. J. Kim, D. Tzamarias, T. Ellenberger, S. C. Harrison, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4513 (1993).
12. H.-M. Wu and D. M. Crothers, *Nature* **308**, 509 (1984).
13. R. P. de Groot, V. Delmas, P. Sassone-Corsi, *Oncogene* **9**, 463 (1994); D. N. Paolella and A. Schepartz, unpublished data.
14. S. S. Zinkel and D. M. Crothers, *Nature* **328**, 178 (1987).
15. T. K. Kerppola and T. Curran, *Science* **254**, 1210 (1991).
16. H.-S. Koo, J. Drak, A. J. Rice, D. M. Crothers, *Biochemistry* **29**, 4227 (1990).
17. M. R. Gartenberg *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6034 (1990).
18. T. K. Kerppola and T. Curran, *Cell* **66**, 317 (1991).
19. ———, *Mol. Cell. Biol.* **13**, 5479 (1993).
20. X-ray crystallographic analysis shows intrinsic curvature toward the major groove within a short DNA fragment containing the central 6 bp of the CRE target sequence (GACGTCA) (H. Rozenberg and Z. Shakked, unpublished data).
21. S. J. Metallo and A. Schepartz, unpublished data.
22. The free energy required to bend DNA smoothly in an elastic model by 20° over 10 bp is <0.5 kcal mol^{-1} (*24*), which would not account for the 50-fold greater stability of the ccc-CRE complex compared to the ccc–AP-1 complex (although localization of the bend to 2 bp would increase the free energy cost to 2.7 kcal mol^{-1}). The binding of the AP-1 sequence by ccc required a greater minor groove bend than did the binding of the CRE sequence (Fig. 3D), which suggests that a more demanding structural rearrangement of the AP-1 sequence is required whose energy cost is not calculated accurately from the change in bend angle. One may draw an analogy between this result and the imperfect correlation between binding strength and extent of bending observed in the CAP system [D. D. Dalma-Weiszhausz, M. R. Gartenberg, D. M. Crothers, *Nucleic Acids Res.* **19**, 611 (1991)].
23. Phasing analysis of the CRE and AP-1 test fragments bound to the chimeric peptide cgg each showed evidence of protein-induced curvature toward the minor groove, mirroring the results obtained with ccc [D. N. Paolella and A. Schepartz, unpublished data].
24. H.-N. Liu-Johnson *et al.*, *Cell* **47**, 995 (1986).
25. T. A. Steitz, *Q. Rev. Biophys.* **23**, 205 (1990).
26. J. Kim *et al.*, *Gene* **85**, 15 (1989).
27. Peptides were synthesized with the use of solid-phase methods, purified by high-performance liquid chromatography, and characterized by amino acid analysis and electrospray ionization mass spectrometry at the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University. Peptide concentrations were determined by amino acid analysis in triplicate with a Beckman 73000 Analyzer with norleucine and homoserine as internal standards.
28. I. Hope and K. Struhl, *Cell* **43**, 177 (1985).
29. T. Maekawa, *EMBO J.* **8**, 2023 (1989).
30. The synthetic oligonucleotides CTAGAGAGATGACGTCATCTCGC (CRE; binding site indicated by underlining) or CTAGAGAGATGACTCATCTCGC (AP-1) were cloned into pBEND2 (*26*) between the unique Xba I and Sal I restriction sites to generate plasmids pDP-CRE and pDP-AP-1, respectively. Plasmids containing the proper insert were identified after transformation into competent *Escherichia coli* GM2929 cells by screening transformants for loss of the Sal I site; plasmids were then sequenced (CircumVent Kit, New England Biolabs) to ensure the integrity of the inserted sequence. Plasmids pDP-CRE and pDP-AP-1 were digested separately with each of the seven restriction enzymes shown in Fig. 2A, and the resulting fragments were purified by native gel electrophoresis to generate two sets of oligonucleotides (CRE_{a-g} and $AP\text{-}1_{a-g}$) that differed only in the position of the CRE or AP-1 site relative to the fragment ends. Oligonucleotides were 5′ labeled with [γ-^{32}P] adenosine triphosphate with the use of T4 polynucleotide kinase.
31. M. Fried and D. M. Crothers, *Nucleic Acids Res.* **9**, 6505 (1981); M. M. Garner and A. Revzin, *ibid.*, p. 3047.
32. Binding reactions were performed by incubation of the indicated peptide (10 to 40 nM monomer) with 10 to 100 pM labeled, double-stranded DNA in a final reaction mixture containing phosphate-buffered saline buffer [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, and 0.05% NP-40 for 30 min at 4°C. Free and peptide-bound DNA were resolved by nondenaturing 8% (32:1) polyacrylamide gel electrophoresis in TG buffer [18.75 mM tris-OH and 162.25 mM glycine (pH 8.9)] at 4°C at 12 V cm^{-1}. Radioactivity was quantified with a Betascope 603 Blot Analyzer (Betagen Corp., Waltham, MA) and by autoradiography.
33. We generated plasmids pDP-AP-1-21, -23, -26, -28, and -30 by cloning the AP-1 site (underlined) containing sequences CTAGAGAGATGACTCATCTGCAAAAACGGGCAAAAACGGGCAAAAACGC, CTAGAGAGATGACTCATCTCTGCAAAAACGGGCAAAAACGGGCAAAAACG, CTAGAGAGATGACTCATCTCGCCTGCAAAAACGGGCAAAAAACGGGCAAAAACGC, CTAGAGAGATGACTCCATCTCGCACCTGCAAAAACGGGCAAAAACGGGCAAAAACGC, or CTAGAGAGATGACTCATCTCGCACTGCTGCAAAAACGGGCAAAAACGGGCAAAAACGC, respectively, between the Xba I and Sal I sites of pBEND2 (*26*). Plasmids pDP-CRE-21, -23, -26, -28, and -30 were prepared in a similar manner, except the AP-1 site was replaced by the CRE (ATGACGTCAT) site. Digestion of each plasmid with Stu I generated the set of oligonucleotides shown.

34. We thank D. Crothers, J. Kahn, T. Shrader, P. Sigler, P. Yang, and N. Zondlo for comments on the manuscript. Supported by the American Cancer Society, with additional support provided by the National Foundation for Cancer Research, Hoffmann–La Roche, Sterling-Winthrop, Pfizer, and Zeneca Pharmaceuticals. A.S. is a David and Lucile Packard Foundation Fellow in Science and Engineering, an Eli Lilly & Co. Grantee in Biochemistry, an NSF Presidential Young Investigator, a Camille and Henry Dreyfus Teacher-Scholar, and an Alfred P. Sloan Foundation Research Fellow. D.N.P. was supported by an NIH Training Grant in Biophysics. C.R.P. is an NSF Predoctoral Fellow.

5 January 1994; accepted 1 April 1994