DNA Targets for Certain bZIP Proteins Distinguished by an Intrinsic Bend

David N. Paolella; C. Rodgers Palmer; Alanna Schepartz


Stable URL:
http://links.jstor.org/sici?sici=0036-8075%2819940520%293%3A264%3A5162%3C1130%3ADTFCBP%3E2.0.CO%3B2-6

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Science is published by American Association for the Advancement of Science. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aaas.html.

Science
©1994 American Association for the Advancement of Science

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR
DNA Targets for Certain bZIP Proteins Distinguished by an Intrinsic Bend

David N. Paoella, C. Rodgers Palmer, Alanna Schepert

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 sequences.

The bZIP family of eukaryotic transcription factors contains (i) a short, helical basic segment whose residues participate in DNA contact, (ii) a zipper segment responsible for protein dimerization ([i], 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 helices that interact with each other along the length of the zipper segment to form a parallel coiled coil (4). These 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 pseudo-symmetric 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 36° (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 ggcc and cacc, 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 configurational distortions near

D. N. Paoella and A. Schepert, Department of Chemistry, Training Program in Biophysics, Yale University, New Haven, CT 06521, USA.
C. R. Palmer, Department of Molecular Biophysics and Biomechanics, Yale University, New Haven, CT 06511, USA.

©1994 by the National Academy of Sciences
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 fragment 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 25 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 apical arms. 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 \( \rho \) as the ratio of the normalized mobilities for the ggg-CRE complexes divided by the corresponding free CRE mobilities \( \mu_{\text{norm}}/\mu_{\text{free}} \). The value of \( \rho \) was within 2% of unity and independent of the spacing 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 13° (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 ATCG 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 conformations 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 9°, an angle similar to that estimated for the free AP-1 sequence (59). 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 major groove (Fig. 3, B and C).
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 unbound 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 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 (pccc, where x is either g or c) and the spacer or zipper segment of either GCN4 or CRE-BPI. In the case of gcc and ggc, the extent of CRE or AP-1 curvature in the complex mimicked the extent of curvature in the free DNA (Fig. 3, B and D). In the case of gcc, which contains only the zipper segment of CRE-BPI, 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-BPI basic segment suggests that residues within the CRE-BPI 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 (22). We conclude that residues within the CRE-BPI 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 energy of the native CRE (filled) 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 (bending) energy gained upon formation of the ccc-CRE 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 bent AP-1 sequence is equal to the (bending) energy required to straighten the intrinsically bent CRE sequence. The (bending) energy gained upon formation of the ccc-AP-1 (bent) complex is less than that gained upon formation of the ccc-CRE (straight) complex.

20. X-ray crystallographic analysis shows intrinsic conformation toward the major groove within a short DNA fragment having the central 6 bp of the CRE target sequence (CARACCA) (H. Ristenpart and Z. Shuker, unpublished data).
22. The free energy required to bend DNA smoothly in an elastic model by 20° over 19 bp is +9.5 kcal mol−1 (34), which would not account for the 5-kcal greater stability of the ccc-CRE complex compared to the ccc-AP-1 complex (although localization of the bend to 10 bp would increase the free energy cost to 27 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. 3A), 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 interferon correlation between binding strength and extent of bending observed in the CRP system (D. G. Dreyfuss, J. M. Schiess, M. R. Gutterson, D. M. Crothers, Nucleic Acids Res. 19, 611 (1991)).
23. Presatting the CRE and AP-1 test fragments bound to the chromatin peptide egg extract showed evidence of protein-induced curvature toward the minor groove, mimicking the results obtained with ccc (D. N. Fadok and A. Scheiz, unpublished data.
27. Peptides were synthesized with the use of solid-phase methods, purified by high-performance liquid chromatography, and characterized by amino acid analysis and mass spectrometry at the W.M. Keck Foundation Bio-technology Resource Laboratory at Yale University. Peptide concentrations were determined by absorbance and amino acid analysis with a 71000 Analyzer with ninhydrin and homogenate at internal standards.
30. The synthetic oligonucleotides CTCAGAGATGAGTCCATGCG (CRE binding site) or CTCAGAGATGAGTCCATGCG (CARACCA) were cloned into pBluescript II to generate the CRE-P1 or CARACCA-P1 vector, respectively. Plasmids containing the proper insert were identified after transformation into competent E. coli DH5α cells by screening transformants for the appearance of blue colonies on X-gal/IPTG plates.
31. First and D. M. Crothers, Nature 305, 659 (1983); M. M. Grann et al. (unpublished observations).