Structure of Homeodomain-Leucine Zipper/DNA Complexes Studied Using Hydroxyl Radical Cleavage of DNA and Methylation Interference†

Adriana E. Tron,‡ Raúl N. Comelli, and Daniel H. Gonzalez*

Cátedra de Biología Celular y Molecular, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, CC 242 Paraje El Pozo, 3000 Santa Fe, Argentina

Received July 8, 2005; Revised Manuscript Received October 17, 2005

ABSTRACT: Homeodomain-leucine zipper (HD-Zip) proteins, unlike most homeodomain proteins, bind a pseudopalindromic DNA sequence as dimers. We have investigated the structure of the DNA complexes formed by two HD-Zip proteins with different nucleotide preferences at the central position of the binding site using footprinting and interference methods. The results indicate that the respective complexes are not symmetric, with the strand bearing a central purine (top strand) showing higher protection around the central region and the bottom strand protected toward the 3′ end. Binding to a sequence with a nonpreferred central base pair produces a decrease in protection in either the top or the bottom strand, depending upon the protein. Modeling studies derived from the complex formed by the monomeric Antennapedia homeodomain with DNA indicate that in the HD-Zip/DNA complex the recognition helix of one of the monomers is displaced within the major groove respective to the other one. This monomer seems to lose contacts with a part of the recognition sequence upon binding to the nonpreferred site. The results show that the structure of the complex formed by HD-Zip proteins with DNA is dependent upon both protein intrinsic characteristics and the nucleotides present at the central position of the recognition sequence.

The homeodomain is a 61-amino acid protein motif found in a group of eukaryotic transcription factors usually involved in regulating developmental processes (*1*−*3*). It folds into a characteristic three-helix structure that is able to specifically interact with DNA. A large group of homeodomains recognize the sequence TAATNN as monomers, although other recognition specificities have been observed. Structural studies of homeodomain−DNA complexes have indicated that specific contacts with DNA are established by residues present in the third helix and in the disordered N-terminal arm (*3*−*6*). Some homeodomain proteins form complexes with each other, which increases the specificity and the affinity of the interaction with DNA (*7*, *8*).

Plant homeodomains constitute a large family of transcription factors that can be divided into different subfamilies (for a review, see ref *9*). The homeodomains of the different subfamilies seem to have diverged before the separation of the branches leading to plants, animals, and fungi (*10*, *11*). One of the subfamilies, termed HD-Zip, seems to be present only in plants. These proteins contain a leucine zipper dimerization motif adjacent to the homeodomain and bind DNA as dimers (*9*, *12*, *13*). The removal of the leucine zipper or the introduction of extra amino acids between the zipper and the homeodomain causes a complete loss of binding, indicating that the relative orientation of the monomers is important for efficient recognition of DNA (*12*, *13*).

There are four classes of HD-Zip proteins, each composed of several members in different plant species (*12*, *14*, *15*). Two of these classes, named HD-Zip I and II, bind a 9-bp dyad symmetric sequence of the type CAAT(N)ATTG, which can be regarded as composed of two partially overlapping TNATTG sequences (*16*). It has been postulated that each monomer interacts with one of these half-sequences in a way that resembles the interaction of monomeric animal homeodomains with DNA. HD-Zip I and II proteins prefer different nucleotides at the central position of the recognition sequence (A/T and G/C, respectively). The specificity for binding at the central position seems to be conferred in part by amino acids 46 and 56 of helix III (Ala and Trp in HD-Zip I and Glu and Thr in HD-Zip II), together with a different orientation of the conserved Arg55 in both proteins, which would be directly responsible for the interaction (*17*).

In the present work, we have employed hydroxyl radical footprinting and interference techniques to analyze the interaction of the sunflower HD-Zip proteins Hahb-4 and Hahb-10, which belong to classes I and II, respectively, with target sites containing A/T or G/C base pairs at the central position. The results are indicative of a different orientation of each homeodomain present in the dimer respective to the TNATTG half-sequence that it binds. The nucleotide present at the central position of each strand in both target sites would be in part responsible for this behavior.

EXPERIMENTAL PROCEDURES

Expression and Purification of Recombinant Proteins. The Hahb-4 and Hahb-10 HD-Zip domains were expressed as fusions with the glutathione *S*-transferase (GST) from *Schistosoma japonicum* as described previously (*13*, *18*).

† Supported by grants from CONICET, ANPCyT, and Universidad Nacional del Litoral. D.H.G. is a member of CONICET; A.E.T. was a fellow of the same Institution.

* To whom correspondence should be addressed. Telephone/Fax: 54-342-4575219. E-mail: dhgonza@fbcb.unl.edu.ar.

‡ Present address: Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115.

10.1021/bi0513150 CCC: $30.25
Published on Web 11/22/2005

Purification by affinity chromatography was carried out essentially as described by Smith and Johnson (*19*), with modifications described by Palena et al. (*20*). Purified proteins (>95% as judged by Coomassie Brilliant Blue staining of denaturing polyacrylamide gels) were used for the assays. Protein amounts were measured as described by Sedmak and Grossberg (*21*) and verified by inspection of the corresponding bands in polyacrylamide gels. The same amounts of Hahb-4 and Hahb-10 yielded similar ratios of free and bound DNA in footprinting experiments with their preferred binding sites, suggesting that the proportion of active protein was similar in both preparations. DNA-binding assays were performed with the proteins fused to GST. Previous experience indicates that the GST moiety does not affect the DNA-binding behavior of these proteins (*13*, *22*−*24*).

Electrophoretic Mobility Shift Assays. Electrophoretic mobility shift assays were performed to separate the free and protein-bound forms of DNA. For this purpose, binding reactions were supplemented with 2.5% Ficoll and immediately loaded onto a running gel (5% acrylamide, 0.08% bis-acrylamide in 0.5× TBE plus 2.5% glycerol; 1× TBE is 90 mM Tris-borate at pH 8.3 and 2 mM EDTA). The gel was run in 0.5× TBE at 30 mA for 1.5 h and subjected to autoradiography to locate the free and bound DNA. Slices of the gel containing these fractions were cut, and DNA was eluted in 1 mL of 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 20 mM EDTA overnight at 4 °C, precipitated, and used for subsequent analysis.

Footprinting Analysis. For the analysis of hydroxyl radical footprinting patterns, double-stranded oligonucleotides containing the HD-Zip I or HD-Zip II binding sites (BS1 or BS2, respectively) with *Bam*HI and *Eco*RI compatible cohesive ends were cloned into pBluescript SK$^-$. From these clones, 44-bp fragments were obtained and labeled in one of their 3′ ends. This was accomplished by PCR using reverse and universal primers, followed by cleavage with either *Hin*dIII or *Xba*I (from the pBluescript polylinker), incubation with the Klenow fragment of DNA polymerase and [α-^{32}P]-dATP, cleavage with the other enzyme, and purification by nondenaturing polyacrylamide gel electrophoresis. Binding of Hahb-4 or Hahb-10 to these oligonucleotides (200 000 cpm) was performed at 20 °C in 15 μL of 20 mM HEPES-KOH (pH 7.5), 50 mM KCl, 1 mM dithiothreitol, 0.5 mM EDTA, 22 ng/μL BSA, 0.5% Triton X-100, 2 mM $MgCl_2$, and 800 ng of HD-Zip protein. After 30 min, the binding reaction was subjected to hydroxyl radical cleavage by the addition of 10.5 μL of 6.6 mM sodium ascorbate, 0.66 mM EDTA (pH 8.0), 0.33 mM $(NH_4)_2Fe(SO_4)_2$, and 0.2% H_2O_2 (*25*). The bound and free oligonucleotides were separated in an electrophoretic mobility shift assay as described above. The corresponding bands were excised from the gel, eluted, and analyzed on denaturing polyacrylamide gels. To determine the position of the footprint within the sequence, a portion of the same fragment was digested with defined restriction enzymes and loaded in the same gel.

Missing Nucleoside Experiments. For missing nucleoside experiments, one-end-labeled oligonucleotides containing the different binding sites were obtained from clones in pBluescript SK$^-$ as described above and subjected to hydroxyl radical cleavage (*25*). Binding of the proteins to the treated oligonucleotide (200 000 cpm) and separation of the free and bound fractions by electrophoretic mobility shift assays were performed as described in the last section. These fractions were excised from the gel, eluted, and analyzed on a denaturing polyacrylamide gel as described above.

Methylation Interference. For methylation interference assays, oligonucleotides labeled in one strand were treated with 0.5 μL of dimethyl sulfate during 10 min in 200 μL of 50 mM sodium cacodylate (pH 8.0) and 1 mM EDTA (*26*). The reaction was stopped by the addition of 300 mM sodium acetate plus 100 mM 2-mercaptoethanol, and DNA was precipitated and used for the binding reactions. Bound and free DNA fractions, eluted as described above, were resuspended in 20 mM potassium phosphate (pH 7.0) and 1 mM EDTA, heated at 90 °C for 10 min, and then transferred to ice. After NaOH was added to a final concentration of 0.1 N, samples were heated again at 90 °C for 5 min, transferred to ice, precipitated, and analyzed on a denaturing polyacrylamide gel.

RESULTS

Hydroxyl Radical Footprinting of Hahb-4/DNA Complexes. We have employed hydroxyl radical footprinting to analyze the binding of the HD-Zip I protein Hahb-4 to different DNA molecules containing binding sites with changes at the central position of the pseudopalindromic CAATNATTG sequence. For this purpose, a 44-bp DNA fragment labeled in one of its 3′ ends was incubated in the presence of a protein containing the Hahb-4 HD-Zip domain and then subjected to hydroxyl radical attack. After that, free and bound DNA were separated and analyzed on a denaturing polyacrylamide gel. The results of Hahb-4 binding to DNA containing the target sequences CAAT(A/T)ATTG (BS1) or CAAT(G/C)ATTG (BS2) are shown in Figure 1 for both strands (we have arbitrarily named top and bottom strands the ones containing a central purine and pyrimidine, respectively). Electrophoretic mobility shift assays performed using different protein concentrations indicated that Hahb-4 displays higher affinity for BS1. The respective dissociation constants (K_{12}) for conversion of dimers bound to DNA into free monomers, calculated as in ref *13*, were 1.6×10^{-14} M^2 and 5.0×10^{-14} M^2 for BS1 and BS2, respectively.

Binding of Hahb-4 to BS1 produces similar, although not identical, protection patterns on both strands, with a stronger intensity covering the half-sequence TNATTG rather than its complement. Different protection intensities are observed, however, at the central position of each strand: while A in the top strand shows strong protection, T in the bottom strand is less protected than surrounding nucleotides. In addition, stronger protection in the top strand is observed in the first four nucleotides of TNATTG, while the last four nucleotides are more protected in the bottom strand.

Footprinting patterns of Hahb-4 binding to BS2 (Figure 1) present subtle but reproducible differences. Here, the protected region in the top strand is less extended than the one observed with BS1, according to the fact that Hahb-4 shows a preference for the latter. It is also noteworthy that the A immediately following the central nucleotide (TN<u>A</u>TTG) is considerably less protected, compared with surrounding nucleotides, in both strands of BS2 than in BS1. The fact that differential protection is observed even in nucleotides that are common to both sequences indicates that the central

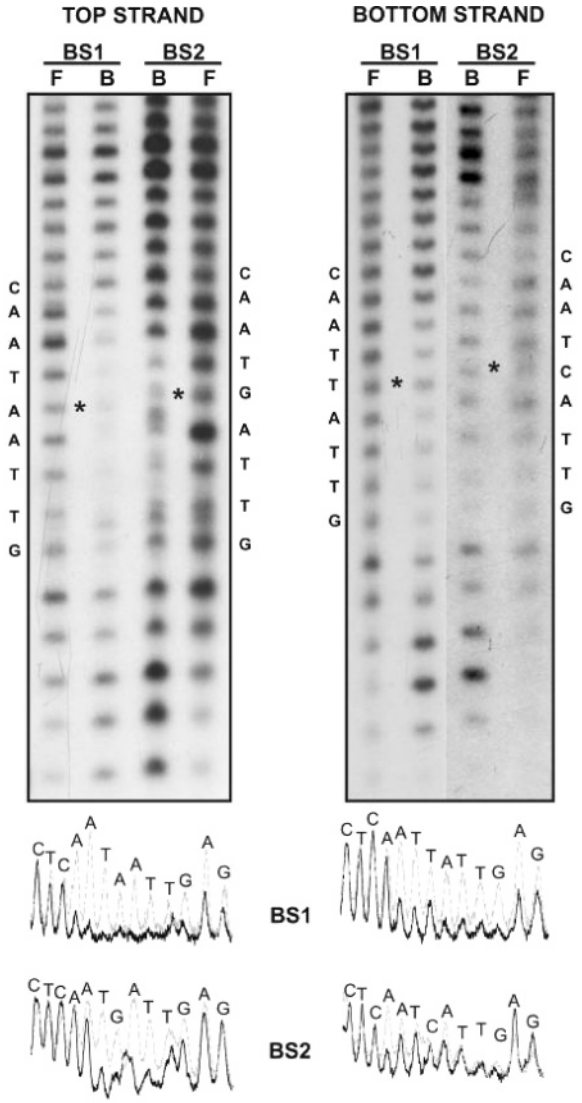

FIGURE 1: Hydroxyl radical footprinting patterns of Hahb-4 complexes with oligonucleotides containing two different binding sites. Hahb-4 was bound to oligonucleotides containing the preferred binding site for either HD-Zip I proteins (BS1) or HD-Zip II proteins (BS2) labeled in the 3′ end of either strand. After binding, DNA was subjected to hydroxyl radical attack. Free (F) and bound (B) DNA were separated and analyzed. A portion of the same fragment digested with defined restriction enzymes was used as a standard to calculate the position of the footprint. Letters beside each panel indicate the DNA sequence (5′ end in the upper part) of the corresponding strand in this region. The strand containing a purine at the central position of the target site was named the top strand in both oligonucleotides. The asterisk indicates the band corresponding to the central position in each case. Below the footprints, densitometric scans of the lanes corresponding to free (broken gray line) and bound (continuous black line) DNA are shown.

base pair influences the positioning of the protein along the entire binding site. It is also interesting to note that binding of Hahb-4 to BS2 protects a nucleotide located in the 5′ region of the bottom strand (CAATCATTG), which shows no or very low protection in BS1. A common feature of the patterns observed with both oligonucleotides is that protection is strongest near the center in the top strand and toward the 3′ end in the bottom strand. This asymmetry may arise from the facts that each monomer establishes different contacts and that only one of them binds the central base pair, closer to the purine nucleotide in both target sites.

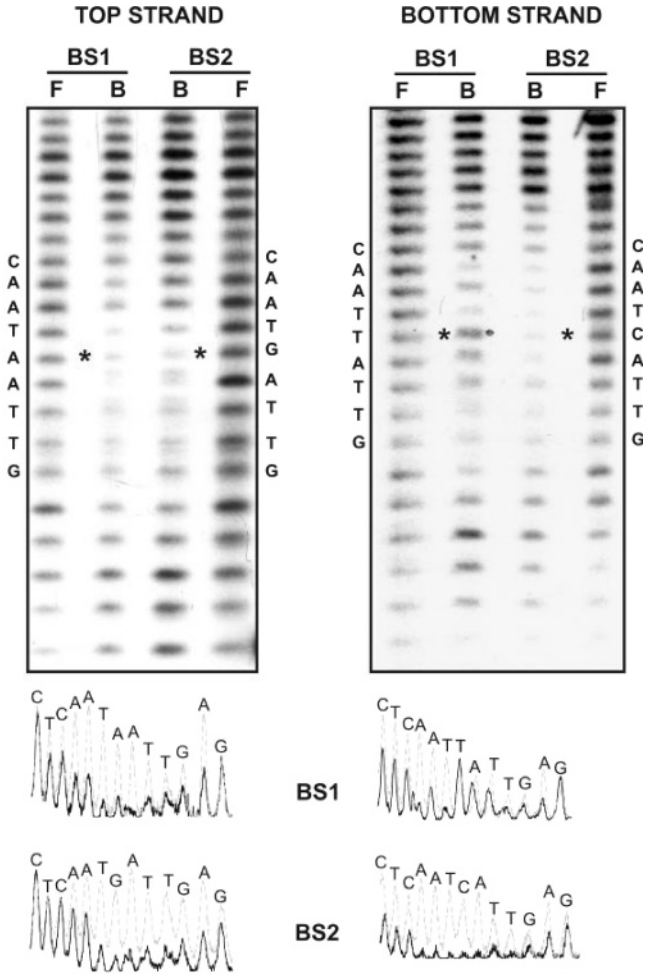

FIGURE 2: Hydroxyl radical footprinting patterns of Hahb-10 complexes with oligonucleotides containing two different binding sites. Hahb-10 was bound to oligonucleotides containing the preferred binding site for either HD-Zip I proteins (BS1) or HD-Zip II proteins (BS2) labeled in the 3′ end of either strand. After binding, DNA was subjected to hydroxyl radical attack. Free (F) and bound (B) DNA were separated and analyzed. A portion of the same fragment digested with defined restriction enzymes was used as a standard to calculate the position of the footprint. Letters beside each panel indicate the DNA sequence (5′ end in the upper part) of the corresponding strand in this region. The strand containing a purine at the central position of the target site was named the top strand in both oligonucleotides. The asterisk indicates the band corresponding to the central position in each case. Below the footprints, densitometric scans of the lanes corresponding to free (broken gray line) and bound (continuous black line) DNA are shown.

Hydroxyl Radical Footprinting of Hahb-10/DNA Complexes. Hydroxyl radical footprinting patterns of Hahb-10/DNA complexes are shown in Figure 2. The dissociation constants for BS1 and BS2 are 2.2×10^{-14} M^2 and 1.2×10^{-14} M^2, respectively. In this case, the patterns observed along the top strand of oligonucleotides containing either BS1 or BS2 are almost identical, with strong protection in the central position of each target sequence. In addition, nucleotides located 3′ to the central position show higher relative protection than nucleotides located 5′ to this site. This pattern is similar to the one observed with the Hahb-4/BS1 complex, indicating that both proteins employ similar mechanisms to bind DNA. Analysis of the bottom strand shows clear differences among both oligonucleotides. Protection in BS1 comprises two regions near the borders of the recogni-

tion sequence, but the central position is not protected. With BS2, protection is observed along the entire recognition sequence and specially in the central region. This may indicate that the preference of Hahb-10 for BS2 is related to a higher number of contacts established mainly with the central region of the bottom strand.

Upon comparing the results obtained with both proteins, it can be observed that each one protects a larger region when interacting with its preferred binding site. For Hahb-4, the main difference between both oligonucleotides is observed in the top strand. In the case of Hahb-10, changes in the bottom strand are more pronounced. It can be assumed that Hahb-10 bound to BS1 adopts a sort of Hahb-4-like binding, establishing weaker contacts with the bottom strand. This may be caused by distortion of the dimer upon binding to a central A/T base pair, resulting in suboptimal orientation of one of the monomers.

Missing Nucleoside Analysis of Hahb-4 Binding to DNA. We have also analyzed the binding of Hahb-4 to oligonucleotides that have been previously subjected to hydroxyl radical attack. This treatment generates a population of DNA molecules with single nucleosides removed at different positions. Nucleosides important for protein binding can then be monitored, because they are under-represented in the bound DNA fraction. Binding of Hahb-4 to BS1 is severely affected by removal of the central nucleoside of the top strand (Figure 3). Two adjacent nucleosides at each side of this position are considerably less important for binding. Strong interference is also observed upon removal of the two last nucleosides of the cognate binding site and one or two adjacent nucleosides, which lie outside the recognition sequence. This indicates that the protein binds to this region, presumably through nonspecific interactions with the sugar−phosphate backbone. The interference pattern in the bottom strand is different, indicating the existence of asymmetric contacts of the dimer with the pseudopalindrome. Interference is observed when nucleosides at the central portion of the binding site are removed, although the pattern is also shifted towards the 3′ end, as with the top strand. Highest interference comprises the sequence TTATT (central position underlined), with nucleosides immediately adjacent to the central T producing a stronger interference than the central one, establishing a pattern with bilateral, rather than dyad symmetry.

The interference pattern of Hahb-4 binding to BS2 shows several differences. In the top strand, removal of T in TGATTG produces a strong relative interference, while a much smaller effect is observed with BS1 at this position. A similar observation can be made for the first T in this sequence (TGATTG). Then, while the central purine seems to be essential for Hahb-4 binding in both target sites, nucleosides adjacent to the central region of the top strand are relatively more important in BS2 than in BS1. A different behavior is observed at the bottom strand. Here, the central and an adjacent nucleoside toward the 3′ end produce a relatively low interference in BS2.

Missing Nucleoside Analysis of Hahb-10 Binding to DNA. The interference patterns observed for Hahb-10 binding to either BS1 or BS2 are identical for the respective top strands (Figure 4). For both oligonucleotides, strong interference is observed at the central and previous position, while the two next nucleosides (TNATTG) produce lower relative interference, as also observed for Hahb-4 binding to BS1. In this

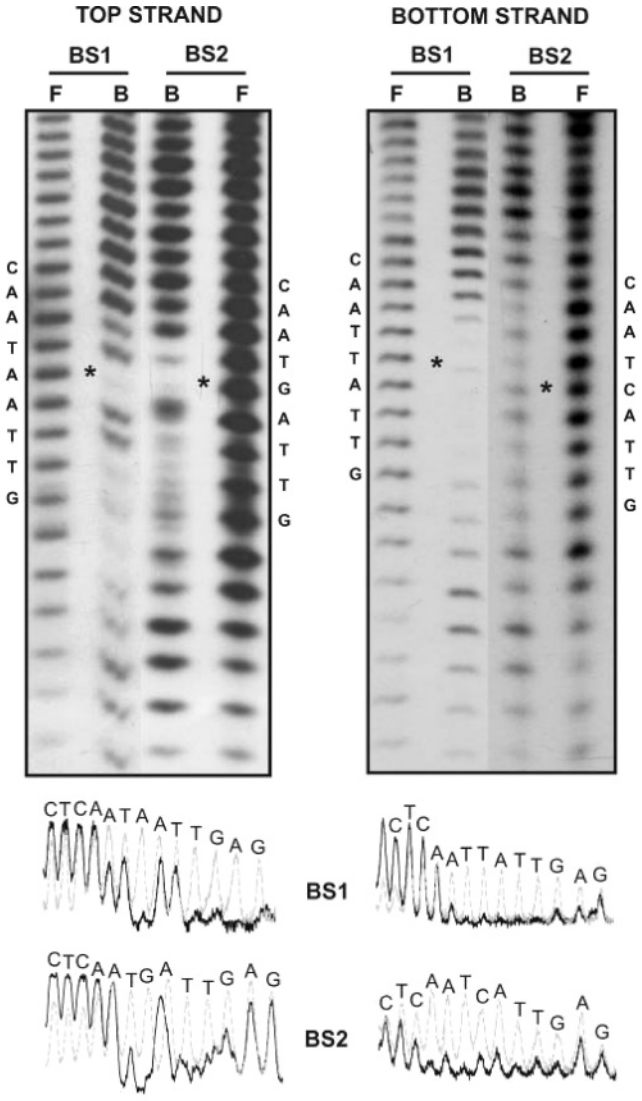

FIGURE 3: Missing nucleoside experiment for the binding of Hahb-4 to DNA. Hahb-4 was bound to oligonucleotides containing the preferred binding site for either HD-Zip I proteins (BS1) or HD-Zip II proteins (BS2) previously labeled in the 3′ end of either strand and subjected to hydroxyl radical attack. After binding, free (F) and bound (B) DNA were separated and analyzed. A portion of the same fragment digested with defined restriction enzymes was used as a standard to calculate the position of the footprint. Letters beside each panel indicate the DNA sequence (5′ end in the upper part) of the corresponding strand in this region. The strand containing a purine at the central position of the target site was named the top strand in both oligonucleotides. The asterisk indicates the band corresponding to the central position in each case. Below the footprints, densitometric scans of the lanes corresponding to free (broken gray line) and bound (continuous black line) DNA are shown.

sense, the interference patterns produced by Hahb-10 with both oligonucleotides resemble the ones produced by Hahb-4 with BS2 and BS1 in the regions located 5′ and 3′, respectively, to the central position.

It is noteworthy that the interference pattern observed for Hahb-10 binding to the bottom strand of BS1 is essentially the same as the one observed for the top strand. Indeed, removal of the central nucleoside of the bottom strand, which shows relatively low protection, strongly interferes with binding (Figure 4). A possible interpretation is that removal of this nucleoside affects the positioning of the central base

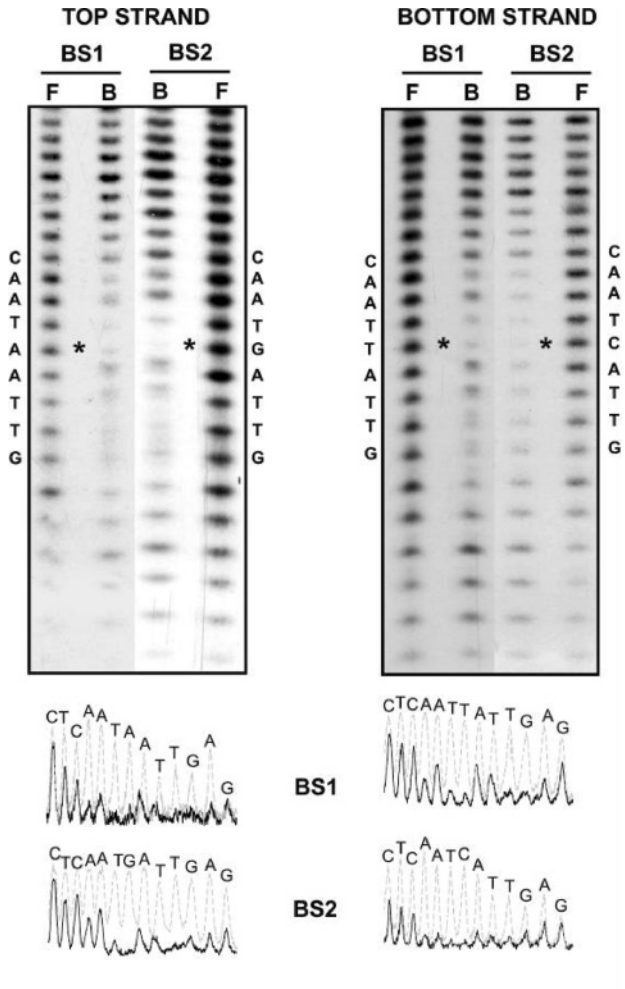

FIGURE 4: Missing nucleoside experiment for the binding of Hahb-10 to DNA. Hahb-10 was bound to oligonucleotides containing the preferred binding site for either HD-Zip I proteins (BS1) or HD-Zip II proteins (BS2) previously labeled in the 3′ end of either strand and subjected to hydroxyl radical attack. After binding, free (F) and bound (B) DNA were separated and analyzed. A portion of the same fragment digested with defined restriction enzymes was used as a standard to calculate the position of the footprint. Letters beside each panel indicate the DNA sequence (5′ end in the upper part) of the corresponding strand in this region. The strand containing a purine at the central position of the target site was named the top strand in both oligonucleotides. The asterisk indicates the band corresponding to the central position in each case. Below the footprints, densitometric scans of the lanes corresponding to free (broken gray line) and bound (continuous black line) DNA are shown.

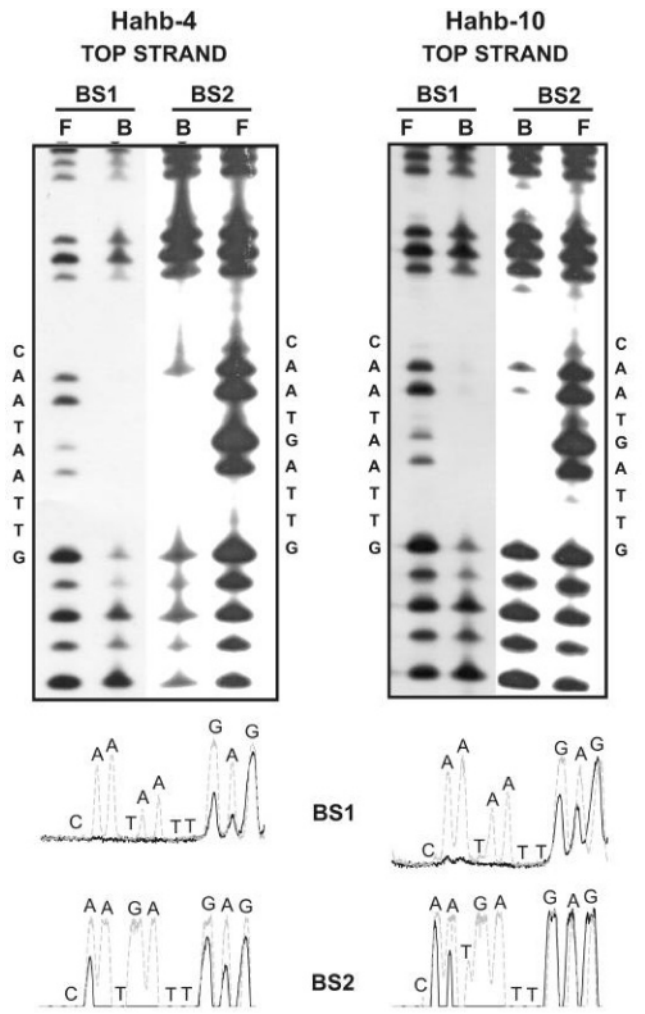

FIGURE 5: Methylation interference of the binding of Hahb-4 and Hahb-10 to DNA. Hahb-4 and Hahb-10 were bound to oligonucleotides containing the preferred binding site for either HD-Zip I proteins (BS1) or HD-Zip II proteins (BS2) previously labeled in the 3′ end of the top strand and subjected to methylation by dimethyl sulfate. After binding, free (F) and bound (B) DNA were separated, cleaved under alkaline conditions, and analyzed. A portion of the same fragment digested with defined restriction enzymes was used as a standard to calculate the position of the footprint. Letters beside each panel indicate the DNA sequence (5′ end in the upper part) in this region. Below the footprints, densitometric scans of the lanes corresponding to free (broken gray line) and bound (continuous black line) DNA are shown.

of the top strand, which is very important for binding. For BS2, the patterns are similar, with the exception that nucleosides located 3′ to the central position show a higher relative interference.

A general comparison of the interference patterns obtained with Hahb-4 and Hahb-10 indicates that contacts with nucleosides 3′ to the central position of the bottom strand are relatively more important for binding of each protein to its preferred oligonucleotide (i.e., Hahb-4 binding to BS1 and Hahb-10 binding to BS2). This indicates that differential binding of both oligonucleotides by these proteins is not only the result of changes in contacts directly established with the central position.

Methylation Interference of Hahb-4 and Hahb-10/DNA Complexes. The importance of nucleotides at specific positions was also investigated through the effect of purine methylation on protein binding. In this way, oligonucleotides with methylated purines at important positions are under-represented in the bound DNA fraction. The results obtained for the top strand are identical with both proteins and indicate that methylation of any purine in the sequence AAT(A/G)ATT severely affects binding (Figure 5). The effect of modification of the first and second A is smaller, however, than the one produced by methylation of the central and adjacent 3′ nucleotides. In addition, methylation of G located in the 3′ end of the target site affects binding to BS1 but not to BS2. Methylation of the bottom strand produces identical results, with the exception that the central nucleotide (a pyrimidine) could not be analyzed in this way.

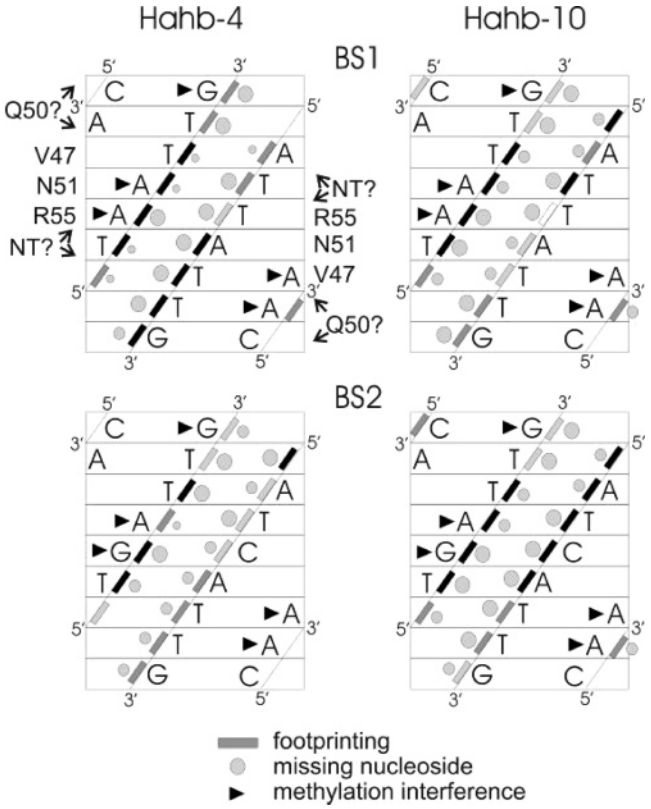

FIGURE 6: Schematic representation of the footprinting and interference experiments. The diagrams show a flat representation of the DNA molecule. Each rectangle represents one base pair and the corresponding major and minor grooves. The top and bottom strands are those that have their 3′ ends in the upper and lower parts, respectively. Bars along diagonal lines (the path of the phosphodiester backbone) indicate the intensity of protection at each position (darker gray means higher relative protection). Circles indicate relative interference at each position after hydroxyl radical cleavage of DNA. Triangles show nucleotides of the top strand that interfere with binding upon methylation. For the Hahb-4/BS1 complex, the putative homeodomain residues that may be involved in interacting with different base pairs, as deduced from general models of the homeodomain/DNA interaction, are shown. NT represents the N-terminal arm of the homeodomain.

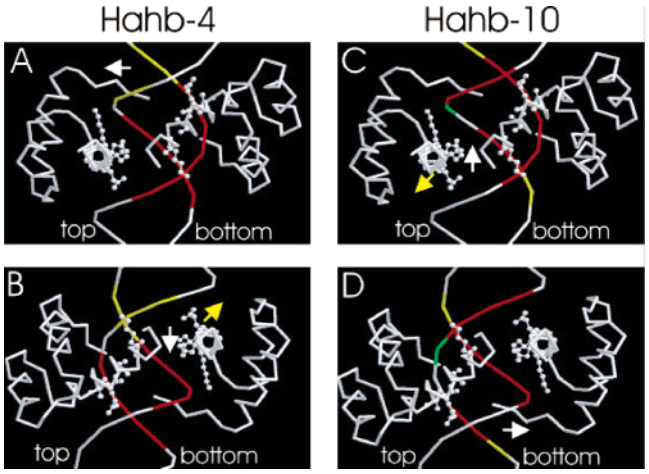

FIGURE 7: Scheme of the HD-Zip domain/DNA complex deduced from the Antennapedia homeodomain/DNA complex. The complex of a single Antennapedia homeodomain with DNA containing the sequence TAATNN was superimposed to each of two TNATTG-binding sites arranged in a 9-bp pseudopalindrome of the type CAATNATTG bound by HD-Zip proteins, thus reflecting the putative position of each monomer of an HD-Zip dimer, using Pdb Viewer (*28*). The images were visualized using RasMol. A and C, on one side, and B and D, on the other, represent views along the axis of helix III of each of the monomers. The DNA regions protected by binding of Hahb-4 (A and B) or Hahb-10 (C and D) to their preferred target sites are shown in color (red for strong protection and yellow for weak protection). The central position of the bottom strand of BS2 in C and D, which is highly protected by Hahb-10, is shown in green for reference. White arrows indicate the changes in the orientation of one of the monomers, respective to the Antennapedia HD, that explain the protection patterns observed for Hahb-4 and Hahb-10 with their preferred binding sites. Yellow arrows indicate changes in relative orientation produced upon binding to the nonpreferred binding site. Note that changes are symmetrical but occur in different monomers (identified by their positioning respective to either the top or bottom strand) of the HD-Zip dimer for Hahb-4 and Hahb-10.

DISCUSSION

Models for the interaction of HD-Zip proteins with DNA have been derived from the structures of complexes formed by animal homeodomains, assuming that each monomer binds to a TAATTG moiety of a pseudopalindromic binding site (*16*). These models have been validated through the use of site-directed mutants and modified binding sites in EMSA experiments (*13*, *16*, *17*, *23*). In the present work, we have made use of footprinting and interference techniques to obtain additional information on the structure of complexes of HD-Zip I and II proteins with DNA. The results obtained after analyzing the binding of Hahb-4 (HD-Zip I) and Hahb-10 (HD-Zip II) to oligonucleotides BS1 and BS2, containing a 9-bp pseudopalindrome composed of two TNATTG half-sites, are summarized in Figure 6. A comparison of the protection patterns for the binding of each protein to its preferred binding site suggests that the proteins adopt a different orientation, with Hahb-4 protecting preferentially the 3′ region of the bottom strand, while Hahb-10 protects mainly the 5′ and central region. Upon binding to the nonpreferred binding site, some contacts seem to be preferentially lost, like the ones with the A next to the central position of the top strand (TGATTG) in the Hahb-4/BS2 complex and those with the central position of the bottom strand in the Hahb-10/BS1 complex.

Figure 7 shows a model of the binding of two homeodomains to a pair of TNATTG half-sites arranged in a 9-bp pseudopalindrome, derived from the complex formed by a single Antennapedia homeodomain with a TAATTG site (*27*). The regions protected in each DNA strand by Hahb-4 binding are indicated in red and yellow for strong and weak protection, respectively (parts A and B of Figure 7). Positions with no or very low protection are shown in white. It can be observed that protected regions do not have the same spatial orientation with respect to each homeodomain. Helix III of one of the homeodomains crosses the major groove, strongly protecting nucleotides at both sides, that is within the ATTG region of the bottom strand and one or two complementary nucleotides of the top strand (Figures 6 and 7A). The other homeodomain protects mainly the central region of the top strand (Figures 6 and 7B). We propose that this would be the result of a different orientation of helix III within the major groove, brought about by the interaction of Arg55 of this monomer with the central A of the top strand (Figure 6). This would bring helix III closer to the top strand relative to the Antennapedia homeodomain (white arrow in Figure 7B). We also propose that T located 5′ to the central A would

be contacted by the N-terminal arm of this monomer (Figure 6), mainly through interactions of positive charges with the sugar−phosphate backbone. It has been previously shown that positive charges within the N-terminal arm are important for HD-Zip protein binding to DNA (*22*). The model of interaction presented in Figure 7 explains the observed effect of the central position on the protection of nucleotides that are identical in both strands.

A similar model in which nucleotides protected by Hahb-10 are represented (parts C and D of Figure 7) shows that the complex formed in this case is much more symmetric. However, certain differences are evident in the relative orientations of the monomers. One of the monomers would be responsible for the protection of the 3′ and the central region of the bottom strand, which is much less protected by Hahb-4 binding to BS1 (Figures 6 and 7C). This would occur, in part, by placement of helix III closer to the bottom strand (white arrow in Figure 7C). The other monomer protects nucleotides at both sides of the major groove, including the central and 3′ region of the top strand and the 5′ region of the bottom strand (Figures 6 and 7D). Binding of Hahb-10 to BS1 is followed by a loss of protection in the central region of the bottom strand (Figure 6). This may be caused by the loss of contacts of one of the monomers with DNA (yellow arrow in Figure 7C). For Hahb-4 binding to BS2, it can be postulated that the monomer that binds the central region of the top strand in BS1 displaces within the major groove, thus protecting 5′ regions of the bottom strand and releasing contacts with the central portion of the top strand (yellow arrow in Figure 7B).

It is also interesting that nucleotides located in the minor groove adjacent to the major groove that is crossed by helix III show different protection depending upon the monomer that is analyzed (Figure 7). Binding through the minor groove has been assigned to contacts established by the N-terminal arm in animal homeodomains. This explains the position of the N-terminal arms in the model presented in Figure 7, which is based on the Antennapedia homeodomain. If protection by HD-Zip proteins in these regions arises from similar interactions, then only one of the N-terminal arms of the dimer seems to establish close contacts with DNA. The remaining N-terminal arm would lie in a different orientation (white arrows in parts A and D of Figure 7), maybe because of steric interference between the monomers. In support of a role of the N-terminal arm in binding to DNA, it has been observed that its removal causes a severe decrease in the binding affinity in the HD-Zip protein Hahb-4 (*22*). In fact, the N-terminal arm and helix I of Hahb-4 can functionally replace the same segment of the animal homeodomain engrailed, at least regarding its capacity to interact with DNA *in vitro* (*24*). The methylation interference assays described here, in addition, support the notion that HD-Zip proteins contact DNA through the minor groove, because adenines are methylated at N3, which faces the minor groove.

Missing nucleoside experiments highlight the importance of contacts established with the central position for binding. Removal of the central nucleoside of the top strand has profound effects on the interaction of both proteins with both target sites (Figure 6). It is noteworthy that A next to the central position of the top strand shows little interference in all interactions. According to models based on animal homeodomains, this A must be contacted by Asn51 of one of the monomers, an essential interaction conserved in almost all homeodomains (*3*). In addition, Sessa et al. (*16*) reported that mutation of this A produces an important decrease in DNA binding by the HD-Zip I protein Athb-1. It can be speculated that the loss of this contact in one of the monomers does not significantly affect binding or that Asn51 of one of the monomers reaccommodates to establish contacts with a different nucleotide.

As a general conclusion, the results described here indicate that HD-Zip proteins form asymmetric complexes with DNA, with helix III of each monomer located in a different orientation respective to the major groove. The asymmetry may arise from contacts with the central position of the target site, which is different in each strand. In addition, HD-Zip I and II proteins may intrinsically present subtle differences in the relative orientations of their monomers, thus allowing them to preferentially recognize different nucleotides at the central position. In regard to this, Sessa et al. (*17*) have observed that residues 46 and 56 of the homeodomain are responsible for this different behavior. They have postulated that these residues may influence the orientation of Arg55, thus producing different recognition specificities. The footprinting results suggest that, in the presence of the nonpreferred binding site, one of the monomers adopts a conformation resembling the one observed with the other protein. This reflects the importance of the binding site in determining the spatial orientation of the dimer. Domain swap experiments using chimeras of Hahb-4 and Hahb-10 will be useful to evaluate this point. The different binding preferences of HD-Zip I and II dimers indicate, however, that intrinsic properties of the proteins that determine the relative orientation of the monomers are also important.

ACKNOWLEDGMENT

We gratefully acknowledge Dr. Raquel Chan for helpful comments during the course of this work.

REFERENCES

1. Gehring, W. J. (1987) Homeo boxes in the study of development, *Science 236*, 1245−1252.
2. Levine, M., and Hoey, T. (1988) Homeobox proteins as sequence-specific transcription factors, *Cell 55*, 537−540.
3. Gehring, W. J., Affolter, M., and Bürglin, T. (1994) Homeodomain proteins, *Annu. Rev. Biochem. 63*, 487−526.
4. Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B., and Pabo, C. O. (1990) Crystal structure of an engrailed homeodomain−DNA complex at 2.8 Å resolution: A framework for understanding homeodomain−DNA interactions, *Cell 63*, 579−590.
5. Otting, G., Qian, Y. Q., Billeter, M., Müller, M., Affolter, M., Gehring, W. J., and Wüthrich, K. (1990) Protein−DNA contacts in the structure of a homeodomain−DNA complex determined by nuclear magnetic resonance spectroscopy in solution, *EMBO J. 9*, 3085−3092.
6. Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D., and Pabo, C. O. (1991) Crystal structure of a MATα2 homeodomain−operator complex suggests a general model for homeodomain−DNA interactions, *Cell 67*, 517−528.
7. Goutte, C., and Johnson, A. D. (1993) Yeast a1 and α2 homeodomain proteins form a DNA-binding activity with properties distinct from those of either protein, *J. Mol. Biol. 233*, 359−371.
8. Knoepfler, P. S., Calvo, K. R., Chen, H., Antonarakis, S. E., and Kamps, M. P. (1997) Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1, *Proc. Natl. Acad. Sci. U.S.A. 94*, 14553−14558.
9. Chan, R. L., Gago, G. M., Palena, C. M., and Gonzalez, D. H. (1998) Homeoboxes in plant development, *Biochim. Biophys. Acta 1442*, 1−19.

10. Bharathan, G., Janssen, B. J., Kellog, E. A., and Sinha, N. (1997) Did homeodomain proteins duplicate before the origin of angiosperms, fungi, and metazoa? *Proc. Natl. Acad. Sci. U.S.A. 94*, 13749−13753.
11. Bürglin, T. R. (1997) Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals, *Nucleic Acids Res. 25*, 4173−4180.
12. Ruberti, I., Sessa, G., Lucchetti, S., and Morelli, G. (1991) A novel class of proteins containing a homeodomain with a closely linked leucine zipper motif, *EMBO J. 10*, 1787−1791.
13. Palena, C. M., Gonzalez, D. H., and Chan, R. L. (1999) A monomer−dimer equilibrium modulates the interaction of the sunflower homeodomain leucine-zipper protein Hahb-4 with DNA, *Biochem J. 341*, 81−87.
14. Sessa, G., Steindler, C., Morelli, G., and Ruberti, I. (1998) The Arabidopsis Athb-8, -9, and -14 genes are members of a small gene family coding for highly related HD-ZIP proteins, *Plant Mol. Biol. 38*, 609−622.
15. Lu, P., Porat, R., Nadeau, J. A., and O'Neill, S. D. (1996) Identification of a meristem L1 layer-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox genes, *Plant Cell 8*, 2155−2168.
16. Sessa, G., Morelli, G., and Ruberti, I. (1993) The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities, *EMBO J. 12*, 3507−3517.
17. Sessa, G., Morelli, G., and Ruberti, I. (1997) DNA-binding specificity on the homeodomain−leucine zipper domain, *J. Mol. Biol. 274*, 303−309.
18. Tron, A. E., Bertoncini, C. W., Chan, R. L., and Gonzalez, D. H. (2002) Redox regulation of plant homeodomain transcription factors, *J. Biol. Chem. 277*, 34800−34807.
19. Smith, D. B., and Johnson, K. S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase, *Gene 67*, 31−40.
20. Palena, C. M., Gonzalez, D. H., Guelman, S., and Chan, R. L. (1998) Expression of sunflower homeodomain containing proteins in *Escherichia coli*: Purification and functional studies, *Protein Expression Purif. 13*, 97−103.
21. Sedmak, J., and Grossberg, S. (1977) A rapid, sensitive, and versatile assay for protein using Coomassie brilliant blue G-250, *Anal. Biochem. 79*, 544−552.
22. Palena, C. M., Tron, A. E., Bertoncini, C. W., Gonzalez, D. H., and Chan, R. L. (2001) Positively charged residues at the N-terminal arm of the homeodomain are required for efficient DNA binding by homeodomain−leucine zipper proteins, *J. Mol. Biol. 308*, 39−47.
23. Tron, A. E., Bertoncini, C. W., Palena, C. M., Chan, R. L., and Gonzalez, D. H. (2001) Combinatorial interactions of two amino acids with a single base pair define target site specificity in plant dimeric homeodomain proteins, *Nucleic Acids Res. 29*, 4866−4872.
24. Tron, A. E., Welchen, E., and Gonzalez, D. H. (2004) Engineering the loop region of a homeodomain−leucine zipper protein promotes efficient binding to a monomeric DNA binding site, *Biochemistry 43*, 15845−15851.
25. Dixon, W. J., Hayes, J. J., Levin, J. R., Weidner, M. F., Dombroski, B. A., and Tullius, T. D. (1991) Hydroxyl radical footprinting, *Methods Enzymol. 208*, 380−413.
26. Wissmann, A., and Hillen, W. (1991) DNA contacts probed by modification protection and interference studies, *Methods Enzymol. 208*, 365−379.
27. Fraenkel, E., and Pabo, C. O. (1998) Comparison of X-ray and NMR structures for the Antennapedia homeodomain−DNA complex, *Nat. Struct. Biol. 5*, 692−697.
28. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling, *Electrophoresis 18*, 2714−2723.

BI0513150