The EMBO Journal vol.12 no.9 pp.3507-3517, 1993

The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities

Giovanna Sessa, Giorgio Morelli[1] and Ida Ruberti[2]

Centro di studio per gli Acidi Nucleici, c/o Dipartimento di Genetica e Biologia Molecolare, Università di Roma La Sapienza, P.le Aldo Moro 5, 00185 Rome and [1]Unità di Nutrizione Sperimentale, Istituto Nazionale della Nutrizione, Via Ardeatina 546, 00178 Rome, Italy
[2]Corresponding author

Communicated by W.J.Gehring

The *Arabidopsis* Athb-1 and -2 proteins are characterized by the presence of a homeodomain (HD) with a closely linked leucine zipper motif (Zip). We have suggested that the HD-Zip motif could, via dimerization of the leucine zippers, recognize dyad-symmetric DNA sequences. Here we report an analysis of the DNA binding properties of the Athb-1 homeodomain-leucine zipper (HD-Zip-1) domain *in vitro*. DNA binding analysis performed using random-sequence DNA templates showed that the HD-Zip-1 domain, but not the Athb-1 HD alone, binds to DNA. The HD-Zip-1 domain recognizes a 9 bp dyad-symmetric sequence [CAAT(A/T)ATTG], as determined by selecting high-affinity binding sites from random-sequence DNA. Gel retardation assays demonstrated that the HD-Zip-1 domain binds to DNA as a dimer. Moreover, the analysis of the DNA binding activity of Athb-1 derivatives indicated that a correct spatial relationship between the HD and the Zip is essential for DNA binding. Finally, we determined that the Athb-2 HD-Zip domain recognizes a distinct 9 bp dyad-symmetric sequence [CAAT(G/C)ATTG]. A model of DNA binding by the HD-Zip proteins is proposed.
Key words: Arabidopsis/HD-Zip motif/homeodomain/leucine zipper/protein-DNA interaction

Introduction

The homeobox (HB) was first recognized as a region of sequence similarity shared by several *Drosophila* genes that play important roles in embryonic development (Gehring, 1987). HB genes have subsequently been identified in many animal species (Scott *et al.*, 1989; Laughon, 1991) including *Hydra*, a primitive metazoan (Schummer *et al.*, 1992). The HB sequence encodes a 60 amino acid (aa) motif, known as the homeodomain (HD). *In vitro* DNA binding studies (Desplan *et al.*, 1988; Hoey and Levine, 1988; Muller *et al.*, 1988) and structural studies (Qian *et al.*, 1989; Kissinger *et al.*, 1990; Otting *et al.*, 1990) have revealed that HDs are sequence-specific DNA binding domains that consist of three α-helices and a flexible N-terminal arm. NMR studies of the contacts made in the *Antennapedia* (*Antp*) HD-DNA complex (Otting *et al.*, 1990) and X-ray crystal structure of the *engrailed* (*en*) HD-DNA complex (Kissinger *et al.*, 1990) showed that the HD makes base contacts in both grooves, with helix-3 in the major groove of the DNA and the N-terminal arm in the adjacent minor groove.

The majority of the HD binding sites identified so far share a 5′-TAAT-3′ core sequence that appears to be critical for HD binding (Hayashi and Scott, 1990; Laughon, 1991 and references therein). The importance of the TAAT core is evident in the structures of the *Antp* and *en* HD-DNA complexes. The structural studies established that the conserved amino acids Arg3 and Arg5 make minor groove contacts with the first half of the core (TA), while Ile47 and Asn51 on helix-3 bind the second half of the TAAT core sequence. These studies have also shown that residue 50, which is a major determinant of DNA binding specificity (Hanes and Brent, 1989; Treisman *et al.*, 1989; Percival-Smith *et al.*, 1990; Schier and Gehring, 1992), makes contacts with bases on the 3′ side of the TAAT sequence. NMR studies of the *Antp* HD-DNA complex have shown that a sixth residue, Met54, contacts DNA. The *en* HD contains at position 54 an alanine, which does not extend far enough for base contact.

The crystal structure of the MAT α2 HD-DNA complex has recently been determined and is very similar to the *en* HD-DNA complex, despite there being only 27% sequence identity between the two HDs (Wolberger *et al.*, 1991). As observed with the *Antp* and *en* HD-DNA complexes, the N-terminal arm makes contacts in the minor groove and helix-3 makes an extensive set of contacts in the major groove. However, *en* and α2 use slightly different regions of the N-terminal arm to make minor groove contacts, suggesting that differences in the amino acid sequences and in the orientation of the N-terminal arm may also play a role in determining binding specificity.

The *en* and α2 HDs also make an extensive set of DNA backbone contacts that appear to play a critical role in positioning the HD on the binding site. Most of these residues are highly conserved within members of different HD classes (Scott *et al.*, 1989) and thus are expected to provide crucial reference points to dock the recognition helix against the major groove of the DNA in other HD-DNA complexes (Wolberger *et al.*, 1991). However, the importance of HD-DNA backbone contacts in contributing also to a higher diversity of DNA binding specificity among HD proteins has been recently suggested by Furukubo-Tokunaga *et al.* (1992). They have found that the two *Antp* class-specific DNA backbone contacting amino acids Arg28 and Arg43, are important determinants of the function of the *fushi tarazu* (*ftz*) protein both *in vivo* and *in vitro*.

Recently, HB genes have been identified in two plant species, maize (Vollbrecht *et al.*, 1991; Bellmann and Werr, 1992) and *Arabidopsis* (Ruberti *et al.*, 1991). Despite the differences in plant and animal development (Poethig, 1990) the discovery of HB genes in plants suggests that fundamental regulatory mechanisms that control development may be shared by all higher eukaryotes. Analysis of the maize *kn1* mutants has shown that ectopic expression of the *knotted* gene results in an alteration of leaf morphology (Hake, 1992).

Therefore it is tempting to speculate that HD proteins in plants are involved in differentiation and/or developmental control as they are in animals.

The *Arabidopsis* Athb-1 and -2 HDs exhibit a strong homology with the helix-3 region and the highly conserved residues of most of the HD sequences. The conservation of these residues has suggested that Athb-1 and -2 HDs may adopt three-dimensional structures similar to those of others. The *Arabidopsis* HB genes also contain a second element that potentially codes for a leucine zipper motif (Zip) located immediately 3′ to the HB. Therefore we have proposed that proteins with a contiguous homeodomain–leucine zipper architecture should be referred to as HD-Zip proteins (Ruberti *et al.*, 1991).

Other members of the *Arabidopsis* HD-Zip class of proteins have been characterized: Athb-3 (Mattsson *et al.*, 1992), HAT 4, 5 and 22 (Schena and Davis, 1992) and Athb-4 (Carabelli *et al.*, 1993). Moreover, the isolation of seven new members of this class of protein has been reported recently (Mattson *et al.*, 1993; Morelli *et al.*, 1993). On the basis of sequence homology, we have grouped the HD-Zip proteins characterized so far into two different families: Athb-1, HAT 5 and Athb-3 in the first family (named HD-ZIP I) and Athb-2, HAT 4, Athb-4 and HAT 22 in the second, HD-ZIP II (Carabelli *et al.*, 1993).

The exact spatial register between the HD and the leucine zipper motif in the two families of HD-Zip proteins is similar to that observed between the DNA binding and the dimerization domains in the b-Zip proteins, another class of transcription factors (Vinson *et al.*, 1989). We have suggested that the HD-Zip proteins might use the dimerization domain to juxtapose a pair of DNA contacting surfaces, each of which fits into half of a dyad-symmetric recognition sequence (Ruberti *et al.*, 1991), in analogy with the DNA binding model proposed for the b-Zip proteins (Vinson *et al.*, 1989; O'Neil *et al.*, 1990).

Here we analyze the DNA binding properties of Athb-1, a member of the HD-ZIP I family. We show that the Athb-1 HD-Zip domain forms a dimeric complex which is able to bind a 9 bp dyad-symmetric sequence. Further, we show that a correct spatial relationship between the HD and the leucine zipper motif is crucial for DNA binding. Finally, we determined that the HD-Zip domain of Athb-2, a member of the HD-ZIP II family, recognizes a distinct 9 bp dyad-symmetric sequence.

Results

Expression of the Athb-1 and Athb-2 HD and HD-Zip domains in Escherichia coli

To express the HD and HD-Zip domains of Athb-1 and Athb-2 in bacteria, we used the pGEX expression system (Smith and Johnson, 1988) (for details, see Materials and methods). Diagrams of the constructs utilized for the expression of the Athb-1 and Athb-2 peptides in *E.coli* are shown in Figure 1A and B (for details, see Materials and methods). These bacterially expressed fusion proteins were soluble and were easily purified by affinity chromatography on glutathione–agarose beads. The integrity and purity of the isolated products were checked by SDS–PAGE (Figure 1C).

The HD-Zip domain of Athb-1 specifically interacts with itself

The presence of a putative leucine zipper motif suggested that Athb-1 may form dimeric complexes (Ruberti *et al.*, 1991). To investigate whether the Athb-1 Zip domain could associate with itself, we took advantage of an affinity chromatography assay developed by Blackwood and Eisenman (1991). In the assay a glutathione S-transferase (GST)-HD-Zip-1 fusion protein, coupled to glutathione–Sepharose beads, was used to test for binding of ^{35}S-labeled HD-Zip peptides, produced by *in vitro* transcription and translation of the corresponding sequences. The labeled *in vitro* translation product (Figure 2, lane 1) was incubated with GST-HD-Zip-1 or GST resin and washed under low stringency conditions, and the bound material was eluted and analyzed by SDS–PAGE. The *in vitro* translated product was retained by the GST-HD-Zip-1 resin (Figure 2, lane 3), but not by GST alone (Figure 2, lane 2).

To test the specificity of the interaction, we determined the ability of the HD-Zip-2 peptide (Figure 2, lane 4), produced by *in vitro* transcription and translation of the corresponding Athb-2 sequence, to bind to GST-HD-Zip-1 resin. No significant amount of HD-Zip-2 was retained either by GST-HD-Zip-1 (Figure 2, lane 6) or GST (Figure 2, lane 5). On the contrary, the HD-Zip-2 peptide is efficiently retained by GST-HD-Zip-2 resin (data not shown).

Taken together, these results reveal a specific interaction of the HD-Zip-1 domain with itself.

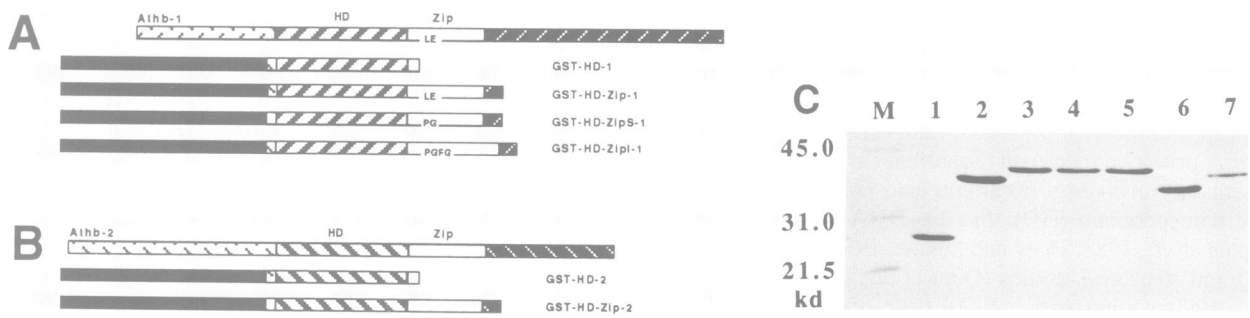

Fig. 1. Structure of GST and GST-fusion proteins. (**A**) Above: the Athb-1 protein (Ruberti *et al.*, 1991) with the positions of the HD (aa 64–124) and the leucine zipper motif (aa 125–160) indicated. Below: the GST-HD-1, GST-HD-Zip-1, GST-HD-ZipS-1 and GST-HD-ZipI-1 fusion proteins. (**B**) Above: the Athb-2 protein (Carabelli *et al.*, 1993) with the positions of the HD (aa 125–185) and the leucine zipper motif (aa 186–220) indicated. Below: the GST-HD-2 and GST-HD-Zip-2 fusion proteins. (**C**) PAGE analysis of purified GST and GST-fusion proteins. Coomassie-stained proteins resolved by SDS–PAGE are shown in lanes 1–7 (GST, lane 1; GST-HD-1, lane 2; GST-HD-Zip-1, lane 3; GST-HD-ZipS-1, lane 4; GST-HD-ZipI-1, lane 5; GST-HD-2, lane 6; GST-HD-Zip-2, lane 7). The numbers indicate the size, in kDa, of protein molecular weight markers (M).

The HD-Zip domain of Athb-1 binds to a 9 bp dyad-symmetric DNA sequence

As a first step towards determining whether the presence of a functional dimerization domain in Athb-1 is required for properly orienting a pair of DNA contacting surfaces, each of which fits into half of a dyad-symmetric sequence, we attempted to identify the binding site of Athb-1. Our strategy for defining the Athb-1 binding site is similar to the procedure recently described by Blackwell and Weintraub (1990). From oligonucleotides in which specific binding site positions are random in sequence, those that are bound are isolated by a chromatography affinity assay, amplified by the polymerase chain reaction, reiteratively re-bound and re-amplified and finally sequenced directly as a pool.

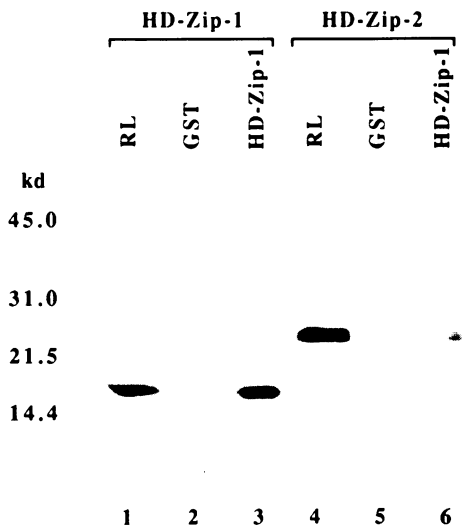

Fig. 2. The HD-Zip domain of Athb-1 specifically interacts with itself. *In vitro* transcription and translation were used to produce peptides labeled with [^{35}S]methionine. After *in vitro* translation, either 0.2 μl programmed reticulocyte lysate (RL) was subjected directly to SDS−PAGE analysis or 10 μl of the lysate was purified on GST or GST-HD-Zip-1 (indicated in figure as HD-Zip-1) Sepharose beads and an aliquot (2 μl) was subjected to SDS−PAGE. Molecular size markers migrated as indicated on the SDS−PAGE analysis.

DNA binding studies on the b-Zip proteins showed that the DNA sequences recognized by these factors consist either of a 10 bp dyad-symmetric sequence in which each half site consists of five bases (like C/EBP, Vinson *et al.*, 1989) or of a 9 bp symmetric sequence in which the central position is contacted by only one of the two monomers of the homodimeric complex (like GCN4, Hill *et al.*, 1986; Oliphant *et al.*, 1989).

For our binding site selection we have designed two different degenerate oligonucleotides, D2 and D1. These oligonucleotides contain two putative half sites giving rise to a 10 bp and a 9 bp sequence, respectively (Figure 3A).

Each half site of the DNA templates used for binding site selection is based on the sequence 5′-TAATNN-3′ (numbered 1−6) which is recognized by the majority of HDs characterized thus far (Laughon, 1991 and references therein). Studies of the *Antp* (Otting *et al.*, 1990) and *en* (Kissinger *et al.*, 1990) HD−DNA complex structures showed that the first two bases of the 5′-TAAT-3′ core sequence are contacted in the minor groove by the N-terminal arm. Arg5 hydrogen bonds to thymine at the first position and Arg3 appears to hydrogen bond to the thymine of the second base pairs (Figure 3A). Positions 3 and 5 in the Athb-1 HD are Glu and Lys, respectively. Therefore, the first two base pairs of the sequence 5′-TAAT-3′ have been randomized in the half site of the DNA templates used for binding site selection. In the *en* structure the positions 3 and 4 of the 5′-TAAT-3′ sequence are contacted in the major groove by Asn51 and Ile47, two of the highly conserved residues on helix 3 (Figure 3A). The position 51 in the Athb-1 HD is Asn and the position 47 is Val. Valine is present at position 47 in many HDs, including *even-skipped*, *paired* and *oct-1*, each of which is capable of high-affinity binding to a TAAT-containing site (Laughon, 1991 and references therein). Therefore, positions 3 and 4 of the

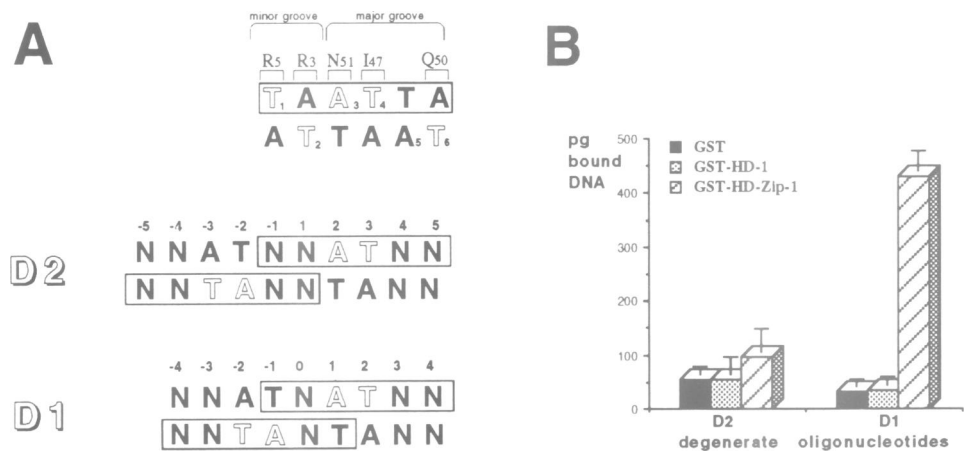

Fig. 3. The HD-Zip domain of Athb-1 binds to specific DNA sequences. (**A**) The core sequence recognized by the *en* HD (Kissinger *et al.*, 1990) and the core sequences of the random-sequence DNA templates D2 and D1, used for binding site selection, are shown. The bases in the sequence recognized by *en* are numbered 1−6; the putative half sites in the 10 bp sequence (D2) are numbered 1−5 (the values are positive for the right half site and negative for the left half site). The central base pair of the 9 bp sequence (D1) is defined as position 0, such that symmetrically disposed base pairs have the same absolute value (positive for positions to the right and negative for positions to the left). The positions of base contact identified in the *en*−DNA complex are shown as brackets at the top. The bases contacted by the *en* HD on each strand of the DNA binding site are indicated by shadowed letters. The TAATTA sequence and the putative half-sites of the DNA templates D2 and D1 are boxed. The core sequences of D2 and D1 are flanked on the 5′ end by primer A and on the 3′ end by primer B (see Materials and methods for complete oligonucleotide sequences). (**B**) Schematic representation of the DNA binding analysis performed using the indicated proteins and the DNA templates D2 and D1. The results shown are the averages of three separate experiments.

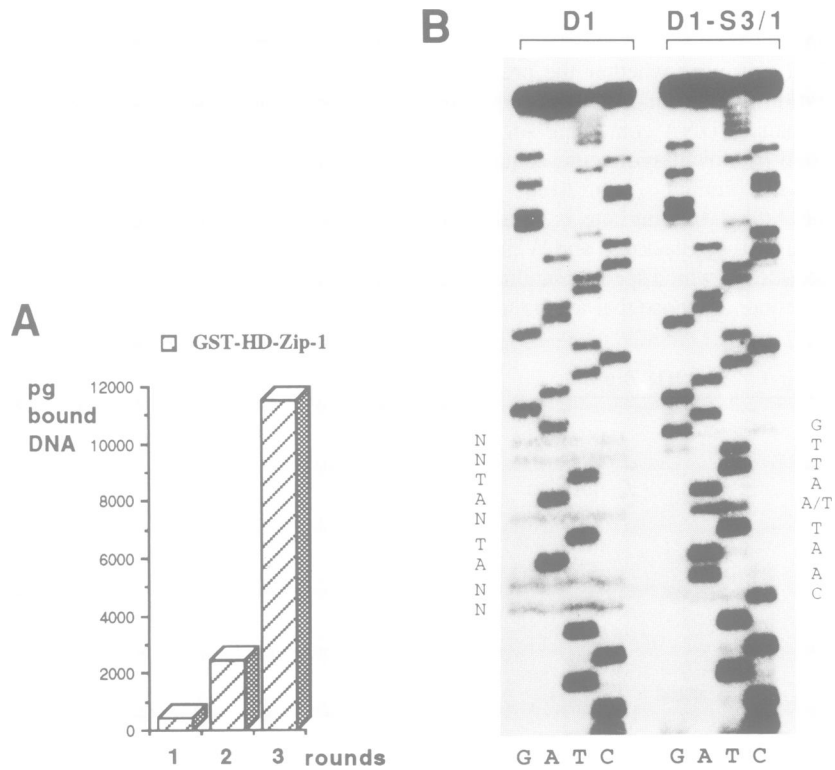

Fig. 4. HD-Zip-1 binding site selection. (**A**) Schematic representation of the HD-Zip-1 binding site selection. (**B**) The sequences of the starting template D1 are compared with those selected for binding by GST-HD-Zip-1. Sequences were determined on amplified DNA by the dideoxy method as described by Blackwell and Weintraub (1990). Those shown were determined with primer B which yields the sequence of the DNA strand shown in the figure. Lanes corresponding to G, A, T and C termination reactions are shown. Putative binding site positions and random sequences of D1 are indicated. Putative binding site positions and experimentally derived sequence preferences of the DNA selected by GST-HD-Zip-1 (D1-S3/1) are also indicated.

sequence 5′-TAAT-3′ were specified in the half site of the DNA templates used for binding site selection (Figure 3A). It should be noted that in the 9 bp sequence, as a consequence of the overlapping of the two half sites, the bases in positions −1 and +1 are specified by the adenines of the opposite strands (Figure 3A).

Attempts to select binding site sequences from the D2 template using *E.coli* extracts which contain either GST-HD-Zip-1 or GST-HD-1 were unsuccessful (Figure 3B). We next used the D1 template (Figure 3A) to determine whether the GST-HD-Zip-1 can interact with DNA. GST-HD-Zip-1 binds to D1 to a significant extent with respect to GST and GST-HD-1 (Figure 3B). The DNA derived from two successive rounds of selection and amplification was bound by the GST-HD-Zip-1 fusion protein to a greater relative extent (Figure 4A). The selected and amplified binding sequences derived from three successive rounds of selection and amplification were determined (Figure 4B). As expected, in the central position of the binding sites were selected two nucleotides, either an A or a T, indicating that it was alternatively recognized by one of the two monomers. Moreover, when confronted with the symmetrical arrangement of random bases in D1, GST-HD-Zip-1 bound with palindromic sequence preferences. A C at position −4 is mirrored by a G at +4, while an A preference at −3 corresponds to a T at +3 (Figure 4B).

To obtain a more accurate definition of the Athb-1 recognition sequence, we performed binding site selection using the DNA template D3, containing a 15 bp random core sequence. Following four rounds of selection and amplification, the DNA was cloned into pBluescript II KS. The sequences of the 25 independently derived, affinity-selected DNA molecules are shown in Figure 5. Strikingly, 18 of these contain a common 7 base dyad-symmetric sequence, AAT(A/T)ATT, and the remaining seven contain sequences differing at only one of the seven positions. This 7 base core sequence was used to align all 25 sequences. By using an alignment in which a pyrimidine residue in the central position (indicated as 0 in the figure) was chosen arbitrarily, the frequencies of each of the bases at 24 positions around the Athb-1 binding element were determined; they are shown in Figure 5. Analysis of the region flanking the core indicates that one more position on each side (±4) has nucleotides that occur at frequencies significantly different from those expected by chance. At positions further from the center (±5 and beyond), the frequencies of nucleotide occurrence are essentially random, with the possible exception of positions −6 and +6. Although the most preferred nucleotides are the same on each side of the binding site, the frequencies of the preferred bases are affected significantly according to their side. In the central nine positions, the majority of deviations from CAATTATTG occurs to the left of the central position, suggesting that Athb-1 probably interacts more avidly with the right half site than with the left half site.

The HD-Zip domain of Athb-1 binds to DNA as a dimer

In order to prove that the CAAT(A/T)ATTG motif is specifically recognized by the GST-HD-Zip-1 fusion protein,

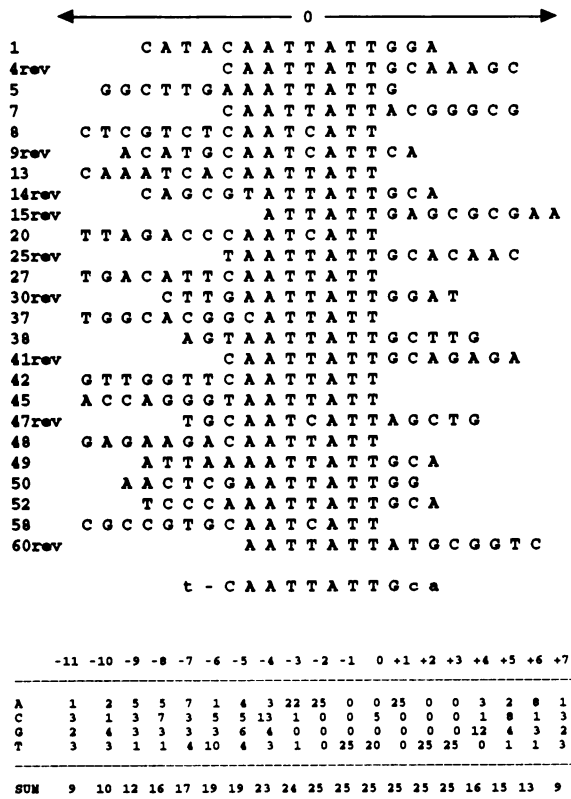

Fig. 5. Tabulation and analysis of the HD-Zip-1 binding site sequences present within affinity-selected DNA molecules. The nucleotide sequences of DNA molecules cloned after the fourth round of selection are shown as aligned by the method described in the text. Only the selected sequence of each DNA is shown. Rev indicates that the reverse complement of the selected sequence is shown. The most common nucleotides for each position between −6 and +6 are indicated. The number of occurrences for the four nucleotides in each position is shown. SUM indicates the total number of nucleotides in each position.

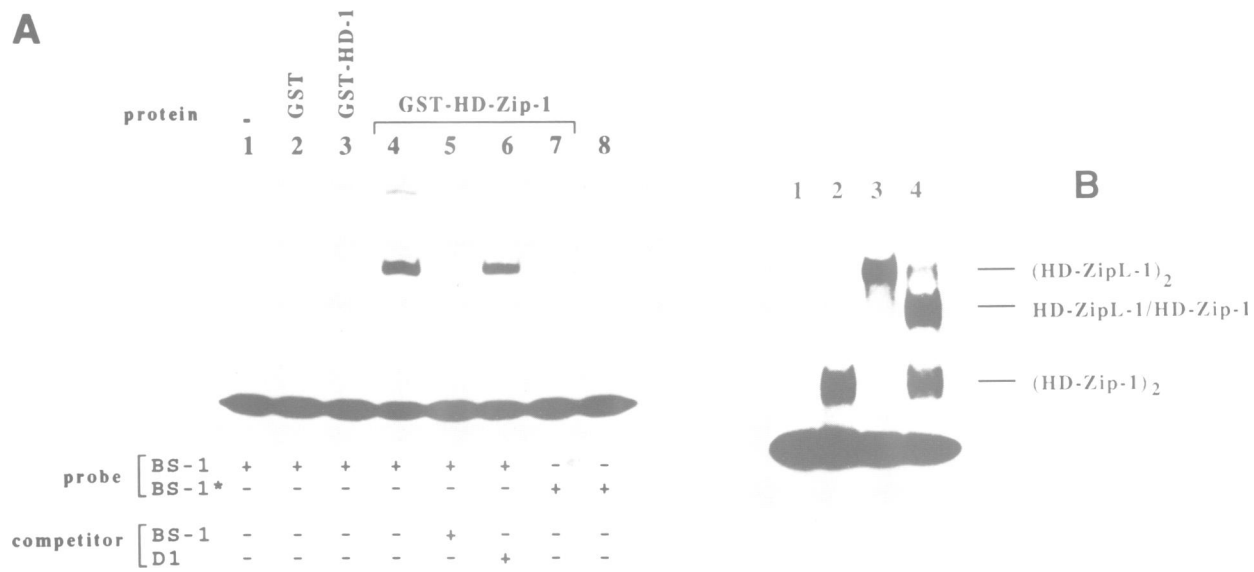

Fig. 6. The HD-Zip-1 domain binds to DNA as a dimer. (**A**) The CAATTATTG motif is specifically bound by the HD-Zip-1 domain. Gel retardation analysis was performed as described in Materials and methods with the oligonucleotide bearing the CAATTATTG motif (BS-1, lanes 1−6) and a derivative containing the CAATTGTTG sequence (BS-1*, lanes 7−8). The retardation assays were performed either with the GST protein (lane 2) or with the GST-HD-1 protein (lane 3) or with the GST-HD-Zip-1 protein (lanes 4−7). To the reaction mixtures shown in lanes 5 and 6 a 100-fold molar excess of unlabeled competitor DNA (BS-1 in lane 5 and D1 in lane 6) was added. (**B**) Mixed-dimer formation by derivatives of Athb-1. The Athb-1 derivatives were assayed for binding to the BS-1 probe individually (lanes 2 and 3) and with the two products cotranslated (lane 4). The control (lacking added RNA) is shown in lane 1.

gel retardation assays (Singh *et al.*, 1986) were performed using the GST-HD-Zip-1 and BS-1, a DNA fragment containing the optimal HD-Zip-1 binding site. The incubation of GST-HD-Zip-1 with BS-1 resulted in a DNA−protein complex of reduced electrophoretic mobility relative to free DNA (Figure 6A, lanes 1 and 4). The formation of the

retarded DNA–protein complex was specific: it was resistant against a 100-fold molar excess of unselected starting DNA template (D1, Figure 6A, lane 6), whereas a 100-fold molar excess of unlabeled BS-1 abolished the formation of the complex (Figure 6A, lane 5). No DNA–protein complex was seen in the assay performed with the labeled DNA and GST alone (Figure 6A, lane 2), indicating that the HD-Zip-1 component of the fusion protein was responsible for DNA binding.

As a first step towards verifying that the HD-Zip domain of Athb-1 binds to DNA as a dimer, the GST-HD-1 protein, which does not have a leucine zipper motif, was tested for its ability to bind to BS-1. In contrast to the GST-HD-Zip-1 protein, the GST-HD-1 does not retard this DNA fragment (Figure 6A, lane 3), indicating that the ability of the GST-HD-Zip-1 protein to recognize DNA is dependent on the presence of the leucine zipper motif. We also performed retardation assays with a derivative of the BS-1 DNA (BS-1*, CAATTGTTG) containing a base substitution (G versus A) in the right half site at the position which should be recognized by Asn51 (see Figure 3A). It has been shown that the effects of mutations at the position contacted by Asn51 are more severe than at any other position of the 6 bp core sequence recognized by the *ftz* HD. Any substitution for adenine at that position resulted in a reduction of at least 25-fold in binding affinity, identifying this base as the single most important component of the *ftz* HD binding site (Florence *et al.*, 1991). No DNA–protein complex formation was seen when a retardation assay was performed with GST-HD-Zip-1 and BS-1* (Figure 6A, lane 7). This observation suggests that the ability of GST-HD-Zip-1 to recognize DNA is dependent on the presence of two intact half sites.

Finally, to demonstrate that the HD-Zip-1 domain binds to DNA as a dimeric complex, the ability of Athb-1 derivatives with different lengths to form mixed dimers was tested by gel retardation assay (Hope and Struhl, 1987) (Figure 6B). When RNAs encoding HD-Zip-1 and HD-ZipL-1 were cotranslated (Figure 6B, lane 4), the products gave an additional protein–DNA complex migrating between the complexes corresponding to HD-Zip-1 (Figure 6B, lane 2) and HD-ZipL-1 (Figure 6B, lane 3).

Taken together these results indicate that the HD-Zip-1 domain binds to DNA as an oligomer and most likely as a dimer.

The leucine zipper in the HD-Zip domain of Athb-1 is required for properly orienting the adjacent helices-3 for DNA binding

It was proposed (Vinson *et al.*, 1989) and recently proved (Pu and Struhl, 1991) that the leucine zipper in the b-Zip proteins correctly and symmetrically positions the two basic regions for specific DNA binding to adjacent half sites. To investigate whether the leucine zipper in the HD-Zip proteins has a similar function, we have analyzed the DNA binding ability of two derivatives of the HD-Zip-1 domain, in which the spatial relationship between the HD and the leucine zipper motif is modified either by substituting the first leucine with a proline or by inserting two amino acids between the two domains (see Figure 1 and Materials and methods). Gel retardation assays performed using the mutant proteins and BS-1 showed that the substitution of the first leucine with a proline (GST-HD-ZipS-1) significantly reduces (by

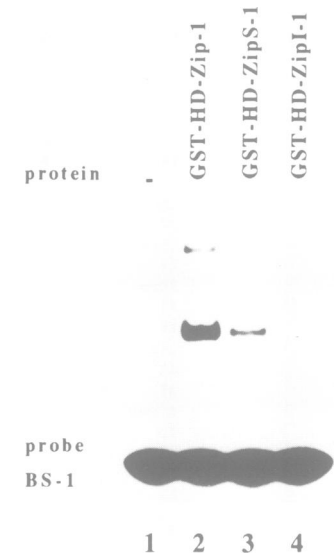

Fig. 7. DNA binding activities of HD-Zip-1 derivatives. Equal amounts of the indicated proteins were incubated with a ^{32}P-labeled DNA fragment containing the optimal HD-Zip-1 binding site (BS-1) and the protein–DNA complexes were electrophoretically separated from unbound ^{32}P-labeled DNA.

~5-fold) the ability of the HD-Zip-1 domain to bind to BS-1 (Figure 7, lane 3), whereas the insertion of two amino acids between the leucine zipper motif and the HD (GST-HD-ZipI-1) abolishes HD-Zip-1 DNA binding activity (Figure 7, lane 4).

To rule out the possibility that both mutations could affect DNA binding activity by interfering with dimerization, we performed affinity chromatography assays with ^{35}S-labeled HD-ZipS-1 and HD-ZipI-1 peptides (see Materials and methods). HD-ZipS-1 and HD-ZipI-1 peptides were specifically retained by GST-HD-ZipS-1 and GST-HD-ZipI-1, respectively (data not shown). The amount of labeled peptides retained by the corresponding fusion proteins was comparable to that of HD-Zip-1 retained by GST-HD-Zip-1 (see Figure 2). The fact that each HD-Zip-1 derivative interacts with itself with an efficiency comparable to that of wild type suggests that the mutations interfere only with the ability of the recognition helices to interact properly with DNA.

The HD-Zip-1 and -2 domains recognize two distinct 9 bp dyad-symmetric DNA sequences

To define the Athb-2 binding site we used the procedure already described for Athb-1 (see Results, section 3). Attempts to select binding site sequences from the D2 template using *E.coli* extracts which contain either GST-HD-Zip-2 or GST-HD-2 were unsuccessful (data not shown). We next used the D1 template to determine whether the GST-HD-Zip-2 protein could interact with DNA. In analogy with the results obtained with the HD-Zip-1 domain, GST-HD-Zip-2 binds to D1 to a significant extent with respect to GST and GST-HD-2 (data not shown). The selected and amplified binding sequences derived from three successive rounds of selection and amplification were determined (Figure 8). In the central position of the binding sites were selected two nucleotides, either a G or a C. The analysis of bases ±3 and ±4 showed that GST-HD-Zip-2, like GST-HD-Zip-1, bound with palindromic sequence preferences. A C or a T

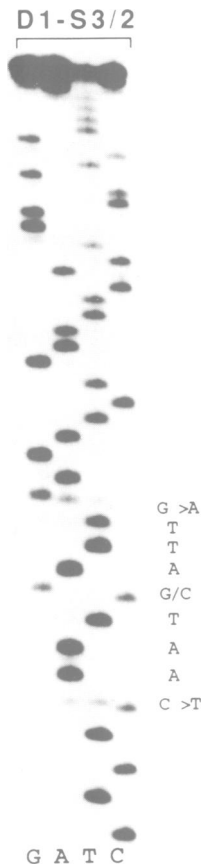

Fig. 8. The HD-Zip domain of Athb-2 recognizes a dyad-symmetric DNA sequence. The DNA selected for binding by GST-HD-Zip-2 was sequenced directly as a pool. Sequences were determined on amplified DNA as described in the legend to Figure 4. Lanes corresponding to G, A, T and C termination reactions are shown. Putative binding site positions and experimentally derived sequence preferences of the DNA selected by GST-HD-Zip-2 (D1-S3/2) are indicated.

at position −4 are mirrored by a G or an A at position +4, and an A preference at position −3 corresponds to a T at position +3 (Figure 8).

The major difference between the HD-Zip-1 (Figure 4B) and -2 (Figure 8) binding sites is observed at the central position, being an A/T and a G/C, respectively.

Discussion

The *Arabidopsis* Athb-1 and -2 proteins are characterized by the presence of a HD with a closely linked leucine zipper motif. We have suggested that the HD-Zip proteins might use the dimerization domain to juxtapose a pair of DNA contacting surfaces, each of which fits into half of a dyad-symmetric recognition sequence (Ruberti *et al.*, 1991). Here we tested several aspects of our model. First, we proved that the HD-Zip domain binds DNA as a dimer. Second, we found that the HD-Zip binding site is a pseudopalindrome of two 5 bp half sites that overlap at a central position. Finally, we showed that the spacing between the zipper and helix-3 is critical for function; in fact, insertion of two amino acids at the beginning of the leucine zipper results in a HD-Zip complex incapable of DNA binding. Our results provide strong experimental support for a model for the DNA binding of the HD-Zip proteins similar to the one proposed by Vinson *et al.* (1989) and O'Neil *et al.* (1990) and recently proved (Ellenberger *et al.*, 1992) for the b-Zip class of transcription factors.

A DNA binding model for the HD-Zip proteins

The putative contacts that each HD-Zip monomer could make at positions 1−4 have been deduced by comparison with the *Antp* (Otting *et al.*, 1990), *en* (Kissinger *et al.*, 1990) and MAT α2 (Wolberger *et al.*, 1991) HD−DNA complexes (Figure 9A and B). In the *Antp* and *en* complexes two more base pairs (positions 1 and 2) are contacted by the N-terminal arm in the minor groove (Figure 9A). These base pairs correspond to positions +1 and 0 in the half site recognized by the HD-Zip-1 and -2 dimeric complexes (Figure 9B). Although further studies will be necessary to evaluate the role of the N-terminal arm in the DNA binding specificity of the HD-Zip complex, for several reasons we favor the hypothesis that the N-terminal arm is not involved in the recognition of the positions 1 and 0. First, the thymine-1 is already specified by the adenine on the opposite strand and therefore its recognition by the N-terminal arm would not contribute to the DNA binding specificity. Moreover, it is unlikely that the GC base pair selected at position 0 by HD-Zip-2 is contacted by the N-terminal arm in the minor groove. In the *en* complex, it appears that the side chain of Arg3 hydrogen bonds with the O2 of the thymine at the equivalent position (Kissinger *et al.*, 1990). Although cytosine residues also provide O2, Florence *et al.* (1991) have shown that a change from thymine to cytosine at this position reduces *ftz* binding 20-fold, probably because the arginine side chain has unfavorable steric or electrostatic interactions with the NH_2 of the guanine. Finally, Florence *et al.* (1991) have observed inhibition for simultaneous binding of two HDs to head-to-head sequences when sites are placed less than four bases apart and suggested that inhibition may be the result of steric hindrances between the N-terminal arms of the HDs. We suggest that the HD-Zip proteins might have lost the ability to make base contacts in the minor groove of the DNA to prevent unfavorable steric interactions between the two N-terminal arms. Supporting this hypothesis is the observation that the Athb HDs are not sufficient for sequence-specific DNA binding activity. In fact, these HDs depend on dimerization for binding even though they contain the majority of the conserved amino acids involved in DNA backbone contacts and the amino acid residues of helix-3 which make specific base contacts in HD complexes (Kissinger *et al.*, 1990; Otting *et al.*, 1990; Wolberger *et al.*, 1991).

The crystal structure of the GCN4 b-Zip element complexed with DNA has recently been determined (Ellenberger *et al.*, 1992). Alignment of the HD-Zip regions with GCN4 revealed that the putative DNA contacting amino acids (Val47, Gln50 and Asn51) of the Athb recognition helices are located, with respect to the leucine zipper motif, in a position equivalent to that occupied by the DNA contacting amino acids of this b-Zip protein (Ala239, Ser242 and Arg243, positions d_{-2}, g_{-2} and a_{-1}, respectively, Figure 9D). Moreover, most of the amino acids contacting the DNA backbone, located in a similar position along the recognition helices of *en*, MAT α2 and GCN4 (positions f_{-3}, a_{-2}, d_{-2}, c_{-1} and g_{-1}, Figure 9D), are also present in the HD-Zip proteins. This array of DNA-contacting amino acids suggests that the b-Zip and the HD-Zip complexes

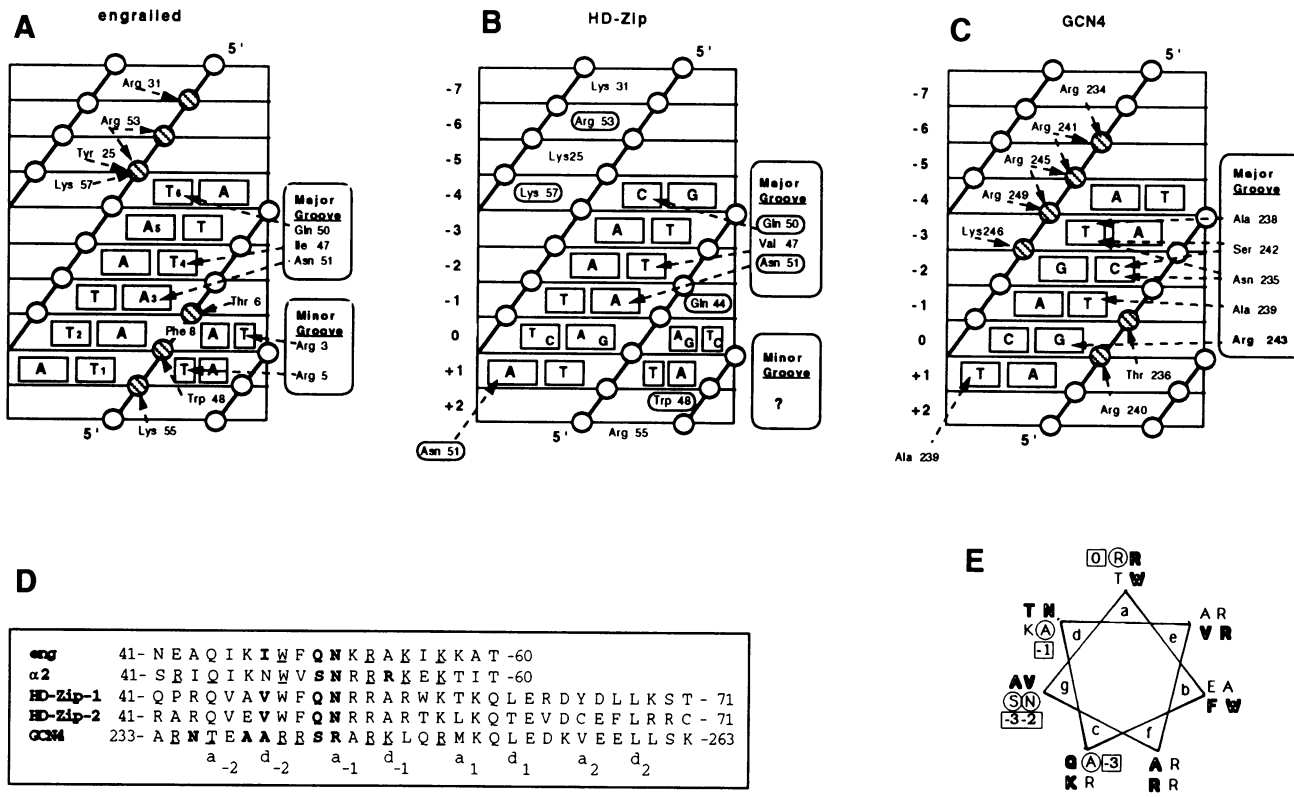

Fig. 9. Comparison of the HD-Zip−DNA complexes with the *en* HD−DNA complex and the GCN4 b-Zip−DNA complex. (**A**) Diagram of contacts made by the *en* HD with its site (adapted from Kissinger *et al.*, 1990). (**B**) Diagram of the putative contacts made by the Athb-1 and -2 HD-Zip domains with their respective sites. In the central position (0) are indicated both T and C which are recognized by HD-Zip-1 and -2, respectively (pyrimidines were chosen arbitrarily, see Results). Residues that are identical in the *en* HD and the two HD-Zip domains are encircled. Gln44, which makes a phosphate contact in the α2 complex, is also indicated. All the amino acids indicated, with the exception of the amino acid at position 25 (Lys and Thr in HD-Zip-1 and -2, respectively), are conserved between HD-Zip-1 and -2. A question mark indicates the amino acids of the N-terminal arm which could be potentially involved in minor groove contacts. (**C**) Diagram of contacts made by the GCN4 b-Zip domain with its site [adapted from Ellenberger *et al.* (1992)]. (**D**) Comparison of the amino acid sequences of HD-Zip-1 and -2 with those of the *en* HD, MAT α2 HD (Wolberger *et al.*, 1991) and GCN4 b-Zip motif. The amino acids of the *en* HD, α2 HD and b-Zip motif that make base contacts are indicated in bold, while those that make DNA backbone contacts are underlined. The putative base-contacting amino acids of the HD-Zip-1 and -2 domains are indicated in bold. The sequence of the GCN4 b-Zip element is shown as a heptad repeat (a−g) with positions a and d corresponding to residues of the dimer interface of the coiled coil. (**E**) Helical wheel analysis of the recognition helix of Athb-1 and the basic region α helix of GCN4. The most N-terminal residue included in the analysis is Ala233 of GCN4, and is placed at position e of the idealized helix. A total of 14 amino acids of GCN4 are included in the analysis, with the most C-terminal residue being lysine 256. 14 amino acids of the Athb-1 recognition helix are displayed on the same schematic α-helix and are indicated by shadowed letters. The most N-terminal residue included is Val45 so that the first putative contacting amino acid, Val47, is located in the same spatial orientation of GCN4 Asn235. The GCN4 contacting amino acids are encircled. The base positions contacted by GCN4 are boxed.

might interact similarly with DNA. Comparison of the putative contacts in the HD-Zip−DNA complex with those in GCN4 b-Zip−DNA complex (Ellenberger *et al.*, 1992) revealed that Asn51 of the HD-Zip and Arg243 of the b-Zip, located in the equivalent position a_{-1}, should contact a different position of their respective binding sites (bp 1 and 0, respectively) (Figure 9B and C). In order to penetrate deeper in the major groove of DNA and make specific contact with base 1, Asn51 on each helix-3 has to be rotated counterclockwise (~50°) with respect to the corresponding Arg243 residue on the GCN4 basic region α helix. This implies that the two recognition helices in the HD-Zip−DNA complex might diverge from each other more than what has been observed for the basic regions of the GCN4 complex. Superimposition of the recognition helices such that Asn51 and Ala239 are now located in the same spatial orientation, results in a consequent alignment of the other base-contacting amino acids (Val47 with Asn235, Gln50 with Ala238 and Ala54 with Ser242) as well as of most of the DNA backbone-contacting amino acids (Trp48 with Thr236, Arg52 with Arg240, Arg53 with Arg241, Lys57 with Arg245) (Figure 9E). Moreover, Arg243, which contacts the central base pair in the b-Zip−DNA complex is now aligned with Arg55 (Figure 9E). We propose that this residue is responsible for specific contact with base pair 0 in the HD-Zip−DNA complex. Local distortion generated by the amino acid residues of the putative fork region, could bring the guanidinium group of Arg55 into an optimal orientation to hydrogen bond to the O4 of thymine-0 in the HD-Zip-1−DNA complex and either the O6 or the N7 of guanine-0 in the HD-Zip-2−DNA complex.

It should be emphasized that the assignments of the contacts in the HD-Zip−DNA complexes are speculative and based on the DNA contacts made by *Antp*, *en*, α2 and GCN4 proteins. Mutational studies will be required to judge the functional relevance of our predictions.

Conclusions

HD-Zip proteins interact with DNA recognition elements in a fundamentally different fashion from the classic HD

Table I. DNA sequence of the oligonucleotides

Oligonucleotide	Sequence	
Athb-1 5D	5′-CCCGGGATCCGCCATGGATCTCTACGACGATG-3′	(positions 489–505 of the cDNA)
Athb-1 3Sma	5′-GAACCCGGGCTGTTTTGTTTTCCAACGAGC-3′	(682–702)
Athb-1 A3D	5′-CCCGGGAATTCGAAGATCGTAGTCTCTCTC-3′	(706–724)
Athb-1 5Sma	5′-GAACCCGGGAGAGACTACGATCTTCTC-3′	(709–726)
Athb-1 3D	5′-CCCGGGAATTCGACCAGGTGGTTCATTAG-3′	(842–859)
Athb-1 B3D	5′-CCCGGATCCTAAGGCCATCCCCAGA-3′	(1133–1149)
Athb-2 5D	5′-CCCGGGATCCGCCATGGGCTCAAGAGGAATCAG-3′	(517–533)
Athb-2 A5D	5′-CCCGGGATCCGCCATGGACGATGAAGATGGTG-3′	(535–550)
Athb-2 A3D	5′-CCCGGGAATTCAGAACTCGCAGTCTACCTCC-3′	(735–754)
Athb-2 B3D	5′-CCCGGGAATTCGGCTCATGTGCATGTAGAA-3′	(844–862)
Athb-2 3D	5′-CCCGGGAATTCCGTCAAAAGTCAAGCCGTG-3′	(994–1012)
Primer A	5′-GTAAAACGACGGCCAGT-3′	
Primer B	5′-GTTTTCCCAGTCACGAC-3′	
D1	5′-GTAAAACGACGGCCAGTGAATTCAGATCTNNATNATNNAGAGGATCCCTCGAGGTCGTGACTGGGAAAAC-3′	
D2	5′-GTAAAACGACGGCCAGTGAATTCAGATCTNNATNNATNNGGATCCCTCGAGGTCGTGACTGGGAAAAC-3′	
D3	5′-GTAAAACGACGGCCAGTGAATTCAGGNNNNNNNNNNNNNNNGGATCCCTCGAGGTCGTGACTGGGAAAAC-3′	
BS-1	5′-GTAAAACGACGGCCAGTGAATTCAGATCTCAATTATTGAGAGGATCCCTCGAGGTCGTGACTGGGAAAAC-3′	
BS-1*	5′-GTAAAACGACGGCCAGTGAATTCAGATCTCAATTGTTGAGAGGATCCCTCGAGGTCGTGACTGGGAAAAC-3′	

proteins, thereby constituting a distinct class of regulatory proteins. The apparent uniqueness of HD-Zip proteins to higher plants suggests that these factors might control developmental pathways that are peculiar to plants. Plant development is distinguished by its plasticity, which arises from a close coupling of developmental responses to environmental stimuli. Of these, light is one of the most important environmental factors influencing the development of plants. In view of this fact, it is remarkable that changes in light quality which profoundly affects plant growth and development also result in a rapid and strong induction of the Athb-2 and -4 expression (Carabelli *et al.*, 1993). This observation suggests that HD-Zip gene products may be an important part of the signal transduction pathway by which light exerts some of its effects on plant growth and development.

What are likely to be the roles of the HD-Zip domain that may confer specific advantages in regulating plant developmental processes? It is evident that the formation of homo or heterocomplexes between members of the HD-Zip families might increase the number of DNA binding or regulatory properties of the HD proteins. In addition, it should be considered that a HD-Zip protein, unable to bind the DNA in the monomeric form, might be turned into a highly specific DNA binding factor by simply promoting the formation of a dimeric complex. This might result in a fast and specific activation of plant cellular responses induced by distinct environmental stimuli.

Materials and methods

Construction of GST fusion proteins

The DNA fragments corresponding to the HD and HD-Zip domains of Athb-1 and Athb-2 were obtained by PCR amplification. The PCR mixture was subjected to a denaturation step at 94°C for 5 min, then to 30 amplification cycles each consisting of a 15 s denaturation step at 94°C, a 15 s annealing step at 60°C and a 1 min elongation step at 72°C. The Athb-1 (Ruberti *et al.*, 1991) and Athb-2 (Carabelli *et al.*, 1993) cDNAs were amplified using primers containing restriction sites for cloning (Athb-1 5D, A3D and 3D, and Athb-2 A5D, A3D and B3D, see Table I). The DNA fragments were ligated in-frame into the *Bam*HI and *Eco*RI sites of the GST expression vector pGEX-2T (Smith and Johnson, 1988). The fusion proteins derived from Athb-1 and -2 were designated GST-HD-1 and GST-HD-Zip-1, and GST-HD-2 and GST-HD-Zip-2, respectively. Two derivatives of the HD-Zip domain of Athb-1, designated HD-ZipS-1 and HD-ZipI-1, were also constructed. HD-ZipS-1, containing two amino acid substitutions at positions d_1 and e_1 of the leucine zipper domain, was constructed by converting the two triplets CTT GAG, corresponding to amino acids Leu and Glu, into a *Sma*I site (CCC GGG, Pro and Gly). The Athb-1 cDNA was amplified using either primers Athb-1 5D and Athb-1 3Sma or primers Athb-1 5Sma and Athb-1 3D (see Table I). The amplified *Sma*I–*Eco*RI fragment was cloned into pGEX-2T; this construct was then used as a vector for the cloning of the *Bam*HI–*Sma*I fragment. The construct containing the *Sma*I–*Eco*RI fragment was used as a vector also for the cloning of the *Bam*HI–*Sma*I amplified fragment restricted only with *Bam*HI. The *Bam*HI–*Eco*RI fragment of this construct corresponds to the HD-ZipI-1 derivative in which the amino acids Leu and Glu are replaced by Pro, Gly, Phe and Gly. Expression constructs were transformed into *E.coli* K12 strain JM109, and purification of GST and GST fusion proteins, from 1.5 ml *E.coli* culture, was performed as previously described (Smith and Johnson, 1988). These fusion proteins were purified with yields ranging between 1.6 and 8 μg/ml of culture.

In vitro transcription and translation

The *Bam*HI–*Eco*RI DNA fragments corresponding to HD-Zip-1, HD-ZipS-1 and HD-ZipI-1 (see the above section) were also cloned into pGEM-4Z. A DNA fragment containing, in addition to the HD-Zip coding region, the region coding for the C-terminal portion of the Athb-1 protein (HD-ZipL-1) was also cloned into pGEM-4Z. The DNA fragment was obtained by PCR amplification of the Athb-1 cDNA with primers Athb-1 5D and Athb-1 B3D (see Table I). For Athb-2, a DNA fragment slightly larger than the one cloned into pGEX-2T, was cloned into pGEM-4Z. The DNA fragment was obtained by PCR amplification of Athb-2 cDNA with primers Athb-2 5D and Athb-2 3D (see Table I). The synthetic oligonucleotides Athb-1 5D and Athb-2 5D include a translational start site. The linearized plasmids were transcribed *in vitro* using T7 RNA polymerase under conditions recommended by the Promega *Protocols and Application Guide*.

RNA was purified, ethanol precipitated and stored at −70°C. For *in vitro* translation, 1−2 μg of RNA was used per 50 μl reaction for 90 min at 30°C using a pretreated rabbit reticulocyte lysate (Promega). To generate radioactively labeled HD-Zip peptides, translations were done using L-[^{35}S]methionine (>800 Ci/mmol; New England Nuclear). The labeled peptides were analyzed by SDS−PAGE under reducing conditions. Gels were fixed, treated with 1 M sodium salicylate for 30 min (Chamberlain, 1979), dried and fluorographed.

Protein dimerization assay

The assay is essentially that described by Blackwood and Eisenman (1991). The *in vitro* translated peptides were assayed immediately after synthesis. An aliquot (0.2 μl) of *in vitro* translated material was subjected directly to SDS−PAGE or the programmed reticulocyte lysates (20 μl) were diluted 20-fold into HND buffer (20 mM HEPES, pH 7.2, 50 mM NaCl, 0.1% NP-40 and 5 mM DTT) with BSA (10 mg/ml). Half of this dilution was incubated with either the GST protein or the GST fusion protein coupled to Sepharose beads [~5 μg of protein adsorbed to 10 μl of glutathione−Sepharose (Pharmacia)] for 1 h at 4°C. The resin was then washed four times with MT-PBS [150 mM NaCl, 16 mM $Na_2HPO_4 \cdot 7H_2O$, 4 mM NaH_2PO_4, (pH 7.3)] (Smith and Johnson, 1988) containing NP-40 (0.1%) at 4°C. The bound peptides were eluted with SDS-containing sample buffer and one-fifth of each sample was subjected to SDS−PAGE and fluorography.

Preparation of E.coli protein extracts

Protein extracts were prepared from the *E.coli* strains that express GST and GST fusion proteins. Overnight cultures of *E.coli* were diluted 1:10 in 5 ml of fresh medium and grown for 1 h at 37°C before adding IPTG to 0.1 mM. After a further 5 h of growth, cells were pelleted and resuspended in 1/10 culture volume of binding buffer (20 mM HEPES pH 7.4, 50 mM KCl, 2 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, 0.5% NP-40, 10% glycerol) that contained 1 mM PMSF. Cells were lysed on ice by mild sonication and then subjected to centrifugation at 10 000 *g* for 5 min at 4°C. The supernatant was aliquoted, quickly frozen in dry ice and stored at −70°C. One aliquot of each protein extract was used to determine the concentration of GST or GST fusion proteins. Protein concentrations were determined by estimating the concentration of Coomassie-stained proteins which were electrophoretically resolved by SDS−PAGE. The level of expression for GST and GST fusion proteins varies. Therefore, to perform DNA binding experiments in which the fused protein−DNA ratio is 1:1 and at the same time the concentration of total *E.coli* proteins is comparable, the protein extracts were diluted with extracts prepared from the *E.coli* JM109 strain.

Oligonucleotides and PCR for binding site selection

All DNA oligonucleotides were synthesized on a Beckman-SM synthesizer. The sequences of the D1, D2 and D3 oligonucleotides [modified from the 70mer utilized by Ekker *et al.* (1991)] used for binding site selection are shown in Table I. N indicates an equal mixture of the four bases during synthesis at that position. Primers A and B (see Table I) were used for amplification of D1, D2 and D3 by cycling 20 times at 94°C (30 s), 58°C (30 s) and 72°C (10 s). Cycling was preceded by 3 min at 94°C and followed by 10 min at 72°C as described by Ekker *et al.* (1991). All PCR amplifications were carried out in a 50 μl reaction volume that contained 200 μM of each dNTP, 0.4 μM of each DNA primer and 0.05 U/μl *Taq* polymerase (Perkin Elmer Cetus) in the standard PCR buffer. Double-stranded DNA templates were labeled to a specific activity of ~2 × 10^8 c.p.m./μg by PCR as described by Blackwell and Weintraub (1990).

Binding site selection

Selection of DNA, specifically bound by the GST fusion proteins, was done by subsequent rounds of DNA binding, isolation of bound DNA by an affinity chromatography assay and PCR amplification. Double-stranded ^{32}P-labeled D1 or D2 DNA (75 ng, 1.5 × 10^5 c.p.m.) was incubated with the *E.coli* protein extract (5 μl), in binding buffer which contained 1 μg poly(dI−dC) (Boehringer Mannheim) as non-specific competitor DNA. The *E.coli* protein extracts utilized for binding contained either GST (~40 ng) or GST-HD-1 (~55 ng) or GST-HD-Zip1 (~65 ng) or GST-HD-2 (~55 ng) or GST-HD-Zip-2 (~60 ng). Incubation was performed in a volume of 20 μl for 1 h at 4°C. After the incubation step, the sample was diluted 5-fold in binding buffer and then mixed at 4°C on a rotating platform with 50 μl 50% glutathione−agarose beads (sulfur linkage, Sigma). After absorption for 10 min, beads were collected by brief centrifugation at 500 *g* and washed three times with 1 ml of binding buffer. Bound DNA was recovered by adding 80 μl of ddH_2O and denaturing the DNA−protein complexes at 100°C for 10 min. The sample was briefly centrifuged and the supernatant was counted to estimate the percent of bound DNA with respect to the total. After addition of 5 μg of tRNA carrier, the supernatant was extracted once with phenol and once with chloroform and then precipitated with ethanol. The pellet was resuspended in ddH_2O (40 μl) and approximately one-tenth of the DNA was amplified by PCR (see the above section). The amplification product was purified on a 14% polyacrylamide gel and then eluted and purified as previously described (Blackwell and Weintraub, 1990). Following three rounds of selection, the nucleotide sequences of the D1-bound and starting template populations were determined using primer B as described by Blackwell and Weintraub (1990).

Binding site selection from the D3 template was performed under the same experimental conditions except that, for the first round of selection, the ^{32}P-labeled DNA was used at a higher specific activity (75 ng, 1.5 × 10^6 c.p.m.). Following four rounds of selection, double-stranded 70mer was digested with *Xho*I and *Eco*RI and ligated into similarly digested pBluescript II KS (Stratagene). 25 single colonies were picked and plasmid DNA was prepared (Quiagen tip 100; Diagen). Sequencing was carried out on double-stranded DNA using Sequenase 2 (USB) according to the manufacturer's instructions. DNA sequences were obtained using the M13(−20) primer.

Gel mobility shift assay

The sequences of the BS-1 and BS-1* DNA fragments used in gel retardation assays are shown in Table I. Double-stranded ^{32}P-labeled DNA (3 ng, 6 × 10^4 c.p.m.) was incubated with *E.coli* protein extract (0.2 μl), in binding buffer which contained 1 μg poly(dI−dC) (Boehringer-Mannheim) as non-specific competitor DNA. The *E.coli* protein extracts utilized for gel retardation assays contained either GST (~1.6 ng) or GST-HD-1 (~2.2 ng) or GST-HD-Zip-1 (~2.6 ng) or GST-HD-ZipS-1 (~2.6 ng) or GST-HD-ZipI-1 (~2.6 ng). Incubation was performed in a volume of 20 μl for 1 h at 4°C. Band shift assays with non-radioactively labeled *in vitro* translated proteins (HD-Zip-1 and HD-ZipL-1) were performed in a similar way. Double-stranded ^{32}P-labeled DNA (see above) was incubated with the cotranslated products (2 μl), in binding buffer containing 0.1 μg poly(dI−dC) and 2 μg of a single stranded oligonucleotide (D3) as non-specific competitor DNA. For binding reactions using a single translational product (1 μl of the *in vitro* reaction) the volume of reticulocyte lysate was equalized with a no-RNA-added control translation. Incubation was performed in a volume of 20 μl for 1 h at 4°C. After the incubation step, the mixture was immediately loaded on to 4.5% polyacrylamide gels (acrylamide:*bis*-acrylamide, 29:1). Electrophoresis was performed in 0.25 × TBE (1 × TBE is 89 mM Tris−borate, 2 mM EDTA; pH 8.3) for 2 h at 14 mA. Prior to loading, gels were prerun for 15 min at 14 mA. Dried gels were subjected to autoradiography at −80°C with intensifying screens. When necessary, bands corresponding to free and bound DNA were excised, placed directly into Aquasol-2 (New England Nuclear) liquid scintillation cocktail and counted on a Beckman LS 5000 TD scintillation counter.

Acknowledgements

We wish to thank N.-H.Chua and P.Fragapane for helpful suggestions, and G.Ruberti for critical reading of this manuscript. We would also like to thank Angelo Di Francesco and Roberto Gargamelli for skilled technical assistance. G.S. was supported by a fellowship of the Fondazione Sigma-Tau. This work was supported, in part, by the Fondazione Istituto Pasteur-Fondazione Cenci Bolognetti, Universitá di Roma La Sapienza (I.R.), by Comitato Nazionale Biotecnologie e Biologia Molecolare, CNR (G.M.) and by Piano Nazionale 'Sviluppo di Tecnologie Avanzate Applicate alle Piante', Ministero Agricoltura e Foreste (I.R.).

References

Bellmann,R. and Werr,W. (1992) *EMBO J.*, **11**, 3367−3374.
Blackwell,T.K. and Weintraub,H. (1990) *Science*, **250**, 1104−1110.
Blackwood,E.M. and Eisenman,R.N. (1991) *Science*, **251**, 1211−1217.
Carabelli,M., Sessa,G., Baima,S., Morelli,G. and Ruberti,I. (1993) *Plant J.*, in press.
Chamberlain,J.P. (1979) *Anal. Biochem.*, **98**, 132−135.
Desplan,C., Theis,J. and O'Farrell,P.H. (1988) *Cell*, **54**, 1081−1090.
Ekker,S.C., Young,K.E., von Kessler,D.P. and Beachy,P.A. (1991) *EMBO J.*, **10**, 1179−1186.
Ellenberger,T.E., Brandl,C.J., Struhl,K. and Harrison,S.C. (1992) *Cell*, **71**, 1223−1237.
Florence,B., Handrow,R. and Laughon,A. (1991) *Mol. Cell. Biol.*, **11**, 3613−3623.

Furukubo-Tokunaga,K., Muller,M., Affolter,M., Pick,L., Kloter,U. and Gehring,W.J. (1992) *Genes Dev.*, **6**, 1082–1096.
Gehring,W.J. (1987) *Science*, **236**, 1245–1252.
Hake,S. (1992) *Trends Genet.*, **8**, 109–114.
Hanes,S.D. and Brent,R. (1989) *Cell*, **57**, 1275–1283.
Hayashi,S. and Scott,M.P. (1990) *Cell*, **63**, 883–894.
Hill,D.E., Hope,I.A., Macke,J.P. and Struhl,K. (1986) *Science*, **234**, 451–457.
Hoey,T. and Levine,M. (1988) *Nature*, **332**, 858–861.
Hope,I.A. and Struhl,K. (1987) *EMBO J.*, **6**, 2781–2784.
Kissinger,C.R., Liu,B., Martin-Blanco,E., Kornberg,T.B. and Pabo,C.O. (1990) *Cell*, **63**, 579–590.
Laughon,A. (1991) *Biochemistry*, **30**, 11357–11367.
Mattsson,J., Soderman,E., Svenson,M., Borkird,C. and Engstrom,P. (1992) *Plant Mol. Biol.*, **18**, 1019–1022.
Mattsson,J., Soderman,E., Svenson,M. and Engstrom,P. (1993) *J. Cell. Biochem.*, Suppl. 17B, 31.
Morelli,G., Baima,S., Carabelli,M., Lucchetti,S., Nobili,F., Sessa,G. and Ruberti,I. (1993) *J. Cell. Biochem.*, Suppl. 17B, 43.
Muller,M., Affolter,M., Leupin,W., Otting,G., Wuthrich,K. and Gehring,W.J. (1988) *EMBO J.*, **7**, 4299–4304.
Oliphant,A.R., Brandl,C.J. and Struhl,K. (1989) *Mol. Cell. Biol.*, **9**, 2944–2949.
O'Neil,K.T., Hoess,R.H. and De Grado,W.F. (1990) *Science*, **249**, 774–778.
Otting,G., Qian,Y.Q., Billeter,M., Muller,M., Affolter,M., Gehring,W.J. and Wuthrich,K. (1990) *EMBO J.*, **9**, 3085–3092.
Percival-Smith,A., Muller,M., Affolter,M. and Gehring,W.J. (1990) *EMBO J.*, **9**, 3967–3974.
Poethig,R.S. (1990) *Science*, **250**, 923–930.
Pu,W.T. and Struhl,K. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 6901–6905.
Qian,Y.Q., Billeter,M., Otting,G., Muller,M., Gehring,W.J. and Wuthrich,K. (1989) *Cell*, **59**, 573–580.
Ruberti,I., Sessa,G., Lucchetti,S. and Morelli,G. (1991) *EMBO J.*, **10**, 1787–1791.
Schena,M. and Davis,R.W. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 3894–3898.
Schier,A.F. and Gehring,W.J. (1992) *Nature*, **356**, 804–807.
Schummer,M., Scheurlen,I., Schaller,C. and Galliot,B. (1992) *EMBO J.*, **11**, 1815–1823.
Scott,M.P., Tamkun,J.W. and Hartzell,G.W.,III (1989) *Biochim. Biophys. Acta Rev. Cancer*, **989**, 25–48.
Singh,H., Sen,R., Baltimore,D. and Sharp,P.A. (1986) *Nature*, **319**, 154–158.
Smith,D.B. and Johnson,K.S. (1988) *Gene*, **67**, 31–40.
Treisman,J., Gonczy,P., Vashishtha,M., Harris,E. and Desplan,C. (1989) *Cell*, **59**, 553–562.
Vinson,C.R., Sigler,P.B. and McKnight,S.L. (1989) *Science*, **246**, 911–916.
Vollbrecht,E., Veit,B., Sinha,N. and Hake,S. (1991) *Nature*, **350**, 241–243.
Wolberger,C., Vershon,A.K., Liu,B., Johnson,A.D. and Pabo,C.O. (1991) *Cell*, **67**, 517–528.

Received on April 20, 1993; revised on June 3, 1993