A monomer–dimer equilibrium modulates the interaction of the sunflower homeodomain leucine-zipper protein Hahb-4 with DNA

Claudia M. PALENA[1], Daniel H. GONZALEZ[1] and Raquel L. CHAN[1,2]

Area Biología Molecular, Facultad de Ciencias Bioquímicas y Farmacéuticas (UNR) and Programa Multidisciplinario de Biología Experimental (PROMUBIE, CONICET), Suipacha 531, 2000 Rosario, Argentina

We have analysed the interaction of the sunflower homeodomain leucine-zipper (Hd-Zip) protein Hahb-4 with DNA. The complete Hd-Zip domain from Hahb-4 was able to select specific sequences from a random oligonucleotide mixture that contained a 9-bp core with four fixed and five degenerate positions. Analysis of the binding of some of the selected sequences suggests that Hahb-4 preferentially binds the dyad-symmetrical sequence CAAT(A/T)ATTG. Single-nucleotide replacements at positions 1, 5 or 9 of this sequence produced a decrease in binding of 2–4-fold. DNA binding as a function of protein concentration was non-hyperbolic. This behaviour could be explained by an equation in which dimer formation is a pre-requisite for DNA binding. A global dissociation constant (K_d) of 1.31×10^{-14} M^2 could be calculated. The removal of the leucine zipper promoted a change in specificity and a decrease in binding affinity ($K_d = 5.03 \times 10^{-5}$ M). Mutation of Phe-20 of the homeodomain into Leu completely abolished DNA binding. The mutant protein, however, was able to inhibit DNA binding by the non-mutant form, presumably through the formation of heterodimers. The analysis of this inhibitory effect at different mutant concentrations allowed the estimation of the K_d for the dimer–monomer equilibrium [about $(2–4) \times 10^{-6}$ M]; from this, a K_d of $3–6 \times 10^{-9}$ M for the dimer–DNA complex could be estimated. The results obtained indicate that the formation of dimers is the main factor influencing the interaction of Hahb-4 with DNA. It is proposed that shifts in a dimer–monomer equilibrium could be used within the cell to modulate the interaction of this protein with target genes.

Key words: binding-site selection, DNA–protein interaction, Hd-Zip protein, plant homeodomain, sunflower.

INTRODUCTION

The homeodomain is a 61-amino acid protein motif found in a group of eukaryotic transcription factors generally involved in regulating developmental processes [1–3]. It folds into a characteristic three-helix structure that is able to specifically interact with DNA. Helices I and II are connected by a loop, whereas helices II and III are separated by a turn that gives this region of the homeodomain a resemblance to prokaryotic helix-turn-helix transcription factors. As monomers, however, most homeodomains are able to bind DNA with high affinity, through interactions made by helix III and a disordered N-terminal arm located beyond helix I (for reviews, see [1–3]).

Plant homeodomains have been discovered relatively recently [4–6], and constitute a large family of transcription factors that can be divided into different subfamilies (for a review, see [7]). The homeodomains of the different subfamilies seem to have diverged before the separation of the evolutionary lineages leading to plants, animals and fungi [8,9]. One of the subfamilies, termed homeodomain leucine-zipper (Hd-Zip) because of the presence of a leucine-zipper motif adjacent to the homeodomain [5,6,10–12], seems to be present only in plants, although it is more related to some animal and fungal proteins than to other plant homeodomains [9]. This suggests that the acquisition of a leucine-zipper motif took place later in evolution. There are two main classes of Hd-Zip proteins, each composed of several members in different plant species [12–15]. The interaction of two of these proteins (Athb-1 and -2, classes I and II, respectively) with DNA has been studied by Sessa et al. [16,17]. These authors have found that both proteins bind as dimers a 9-bp dyad-symmetrical sequence of the type CAAT(A/T)ATTG for Athb-1, and CAAT(C/G)ATTG for Athb-2 [16]. 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 Athb-1; Glu and Thr in Athb-2), together with a different orientation of the conserved Arg-55 in both proteins, although other amino acids are clearly involved [17].

The existence of a large number of related proteins poses the question of how specificity is achieved, especially since it has been suggested that members of the same class may form heterodimers [17]. Thus the study of other Hd-Zip proteins seems important to see how related their DNA-binding properties are. For this purpose, we have chosen a member of the less homogeneous class I, namely Hahb-4 from sunflower. Although Hahb-4 can be assigned to class-I Hd-Zip proteins according to sequence conservation within the homeodomain and leucine-zipper motif, it is clearly the most divergent member of this family described up to now. In fact, several positions known to be conserved in other Hd-Zip proteins show different amino acids in Hahb-4. As an example, position 5, known to be important for DNA binding in other homeodomains [2,18], is occupied by Arg instead of Lys; position 56 is Ser in Hahb-4. This has prompted us to study the interaction of Hahb-4 with different DNA sequences to gain information on the specificity of this interaction and on possible target genes. Our results suggest that all class-I Hd-Zip proteins must have the same specificity of DNA binding. We have also observed that a dimer–monomer equilibrium is established at relatively high protein concentrations, suggesting that Hd-Zip

Abbreviation used: Hd-Zip, homeodomain leucine-zipper.

[1] Present address: Cátedra de Biología, Facultad de Bioquímica y Cs. Biológicas (UNL), CC 242 Pje. El Pozo, 3000 Santa Fe, Argentina.
[2] To whom correspondence should be addressed, at Cátedra de Biología, Facultad de Bioquímica y Cs. Biológicas (UNL), CC 242 Pje. El Pozo, 3000 Santa Fe, Argentina (e-mail raqdan@arnet.com.ar).

© 1999 Biochemical Society

proteins may be present in different oligomeric states, and not only as dimers, *in vivo*.

EXPERIMENTAL

Cloning, expression and purification of recombinant proteins

Hahb-4 coding sequences were amplified and cloned in frame into the *Bam*HI and *Eco*RI sites of the expression vector pGEX-3X [19] as described previously [20]. Amplifications were performed using the same 5′ oligonucleotide (representing sequences adjacent to the N-terminus of the homeodomain) and different 3′ oligonucleotides (5′-GGCGAATTCACATTTCTCAGCACCT-CC-3′ or 5′-GCGGAATTCGCGCGTTATACTCTTGC-3′). The resulting proteins contained coding regions for the entire Hahb-4 homeodomain and leucine-zipper motif, or the homeodomain alone (up to amino acid 67 with respect to the N-terminus of the homeodomain) linked in frame to the 3′ end of the *Schistosoma japonicum* glutathione S-transferase cDNA.

Mutants in the N-terminal region of the homeodomain were constructed by ligating two amplified fragments generated by adjacent oligonucleotides carrying the mutated sequences and a *Bgl*II site at their 5′ ends. To generate a *Bgl*II restriction site, silent changes in codons 12, 13 and 14 of the homeodomain were introduced. Additional changes included either Arg-5 (CGG) to Lys (AAG) or Phe-20 (TTT) to Leu (CTT). All constructions were checked by DNA sequence analysis.

For expression, *Escherichia coli* cells bearing the corresponding plasmids were grown and induced as described previously [20]. Purification and cleavage of the fusion products were carried out essentially as described by Smith and Johnson [19], with modifications described by Palena et al. [20]. For unknown reasons, the presence of truncated products composed mainly of glutathione S-transferase was observed in some preparations (see, for example, Figure 1, lane 1). The results obtained with these preparations were essentially the same as those obtained with preparations where these truncated products were not present.

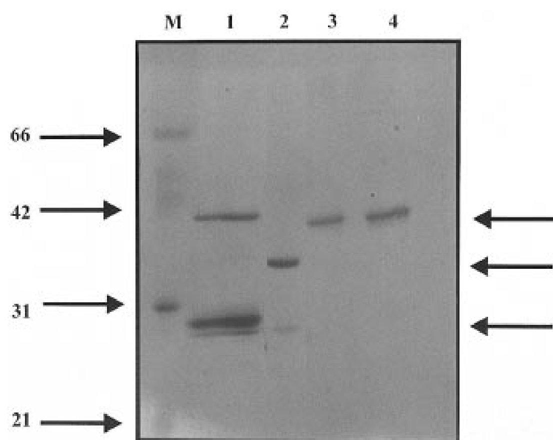


Figure 1 SDS/PAGE analysis of Hahb-4 variants expressed in *E. coli*

Proteins from clones containing different constructs were expressed as fusion proteins and purified as described in the Experimental section. M, molecular-mass markers are indicated on the left (kDa); 1, protein containing the complete Hd-Zip domain; 2, leucine-zipper-deletion derivative; 3, Phe-20 → Leu mutant; and 4, Arg-5 → Lys mutant. The band at about 29 kDa in lane 1 represents a truncated product containing mostly glutathione S-transferase, as suggested by its cleavage by factor X_a; this band was present in some protein preparations and not in others, and its presence did not affect the binding properties of Hahb-4.

Electrophoretic mobility-shift assays

For electrophoretic mobility-shift assays, aliquots of purified proteins were incubated with double-stranded DNA (0.3–0.6 ng, 30000 c.p.m.) labelled with [α-^{32}P]dATP. Reactions (20 μl), containing 20 mM Hepes (pH 7.5), 50 mM KCl, 2 mM $MgCl_2$, 0.5 mM EDTA, 1.0 mM dithiothreitol, 0.5 % Triton X-100, 1 μg of poly(dI-dC) and 10 % glycerol, were incubated for 20 min at room temperature, supplemented with 2.5 % Ficoll and immediately loaded on to a running gel [5 % acrylamide/0.08 % bis-acrylamide in 0.5 × TBE plus 2.5 % glycerol; 1 × TBE is 90 mM Tris/borate (pH 8.3)/2 mM EDTA]. The gel was run in 0.5 × TBE at 20 mA for 2 h and dried prior to autoradiography. For quantitative analysis, poly(dI-dC) was omitted.

Random-oligonucleotide selection

To select DNA molecules specifically bound by Hahb-4, the random-oligonucleotide-selection technique [21] was applied, using a combination of procedures described by Blackwell and Weintraub [22] and Sessa et al. [16]. A 70-mer double-stranded oligonucleotide [16], containing a 9-bp central core with random sequences at 5 positions, was incubated with purified protein as described above. Bound DNA molecules were separated by electrophoretic mobility-shift assays and eluted with 0.5 ml of 0.5 M ammonium acetate/10 mM $MgCl_2$/1 mM EDTA/0.1 % (w/v) SDS [22]. The selected DNA molecules were amplified using oligonucleotides R1 (5′-GTAAAACGACGGCCAGT-3′) and R2 (5′-GTTTTCCCAGTCACGAC-3′). Amplification reactions were performed as follows: 30 cycles of 1 min at 94 °C, 1 min at 53 °C and 1 min at 72 °C. After purification through polyacrylamide gels, the amplified molecules were subjected to new cycles of binding, elution and amplification. Enrichment in sequences bound specifically by Hahb-4 was monitored by competition analysis in electrophoretic mobility-shift assays.

Miscellaneous methods

Double-stranded oligonucleotides were labelled with the Klenow fragment of *E. coli* DNA polymerase I, either by primer extension using oligonucleotides R1 and R2, or by filling in overhanging ends. Total protein was measured as described by Sedmak and Grossberg [23]. When truncated products were present, the actual amount of Hahb-4 was calculated by densitometry of scanned gels. For quantitative analyses, radioactive bands were cut from exposed gels and measured by scintillation counting. Data handling and curve fitting were performed using Sigmaplot software.

RESULTS

Different portions of Hahb-4, consisting of either the complete Hd-Zip domain or the homeodomain alone, were expressed in *E. coli* as fusion proteins with *S. japonicum* glutathione S-transferase, and purified by affinity chromatography as described (Figure 1). In order to identify DNA sequences that were bound specifically by Hahb-4, we used a similar strategy to that used by Sessa et al. [16] to isolate sequences bound by Athb-1 and -2. These two proteins bind a 9-bp dyad-symmetrical sequence of the type ^{1}CAAT(N)ATTG^9 (a central A/T pair is recognized by Athb-1, whereas a C/G pair is bound by Athb-2). Since the nucleotides at positions 3, 4, 6 and 7 are thought to be bound by amino acids 47 and 51 of the homeodomain (Ile/Val and Asn in Hahb-4 and all other Hd-Zip proteins known so far), we have kept these positions constant and randomized the remaining five positions to generate a mixture of 1024 double-stranded oligo-

© 1999 Biochemical Society

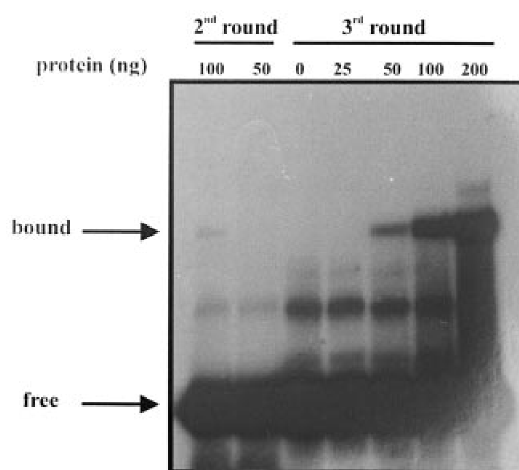


Figure 2 Binding of Hahb-4 to oligonucleotide populations obtained after selection

Different protein concentrations were analysed by electrophoretic mobility-shift assays for binding to oligonucleotide mixtures obtained after two or three rounds of selection.

	NNAT**N**AT**NN**
03	**TA**AT**G**AT**TC**
05	**TA**AT**G**AT**TA**
06	**TA**AT**G**AT**CA**
11	**TA**AT**G**AT**TA**
12	**TA**AT**G**AT**TG**
14	**TA**AT**A**AT**AA**
15	**TA**AT**T**AT**TC**
16	**TA**AT**G**AT**GC**
20	**TA**AT**A**AT**TG**
22	**TA**AT**G**AT**AC**
23	**TA**AT**G**AT**TG**
24	**AA**AT**G**AT**TG**
25	**TA**AT**A**AT**TG**
29	**TA**AT**A**AT**TG**
34	**TA**AT**A**AT**TA**
39	**TA**AT**A**AT**TG**
40	**TA**AT**T**AT**TA**
41	**TA**AT**T**AT**GG**
44	**TA**AT**G**AT**TG**
53	**TA**AT**A**AT**TG**
55	**TA**AT**G**AT**TG**
57	**TA**AT**G**AT**TG**
58	**TA**AT**G**AT**AC**
65	**TA**AT**G**AT**TG**
66	**TA**AT**T**AT**GG**

Position	1	2	5	8	9
A	1	**25**	7	3	6
C	0	0	0	1	5
G	0	0	**14**	3	**14**
T	**24**	0	4	**18**	0

Consensus TAATGATTG

Figure 3 Compilation of the sequences of 25 random clones obtained after cloning the selected oligonucleotide population

The sequences of the central portion of the different clones are indicated in the same orientation with respect to their arms. Clone numbers (arbitrary) are indicated on the left. Variable positions are shown in bold. Below, a table indicating the base frequencies at each position, together with the derived consensus sequence, is shown.

nucleotides from which sequences preferentially bound by Hahb-4 were selected. Figure 2 shows an electrophoretic mobility-shift assay performed with a fusion protein containing the Hd-Zip domain of Hahb-4 and different oligonucleotide mixtures obtained after 2 or 3 rounds of selection by the same protein. A higher proportion of bound DNA was progressively observed, indicating that Hahb-4 was able to select specific DNA sequences from the mixture. After three rounds of selection, no significant improvement could be detected. The bound products were then cloned and 25 randomly picked clones were sequenced. Figure 3 shows a list of the central sequences of these clones, in the same orientations with respect to their arms. Clear deviations from random expectations occurred at all variable positions: positions 1 and 2 were almost exclusively T and A (96 and 100 % respectively); position 5 was G or A in 56 and 28 % of clones, respectively; position 8 was occupied by T in 72 % of the clones; and position 9 was G in 56 % and A in 24 % of clones. Hence, the derived consensus was TAATGATTG (CAATCATTA in the complementary strand). This sequence had two intriguing features when one takes into consideration that Hd-Zip proteins interact with DNA as dimers: first, there was no symmetry at position 1, since T was found in one strand and C in the other strand; second, the central position was biased towards a purine (or a pyrimidine in the complementary strand), whereas either nucleotide of a complementary pair should be found as a result of a symmetric interaction.

To analyse to what extent the derived consensus sequence represented the DNA-binding specificity of Hahb-4, we performed a series of studies using different unique sequences from the isolated clones. Figure 4 (upper panel) shows an electrophoretic mobility-shift assay using different amounts of Hahb-4 and three different oligonucleotides, one of them representing the derived consensus and the other two with single modifications at either position 9 ($G \rightarrow A$, thus forming a symmetrical sequence) or 5 ($G \rightarrow A$). According to this, G was preferred over A at position 9, and A was preferred at position 5.

The question of asymmetry at the central position was tested using the sequences TAATAATTA and TAATTATTA (that is, identical sequences with the arms in reverse orientation). The results obtained indicate a slight preference for A over T at this position, perhaps indicating an influence of the arms in the interaction of Hahb-4 with DNA.

Binding to the asymmetric sequence TAATAATTG was also compared with binding to both symmetric sequences derived from it (TAATAATTA and CAATAATTG). The results (Figure 4, lower panel) indicate that C was preferred at position 1 in both strands. The replacement of one of these C bases by a T reduced binding 2–3-fold, whereas replacing both Cs promoted a further reduction in affinity.

The importance of the four fixed positions (3, 4, 6 and 7) for binding was tested using oligonucleotides with symmetrical modifications at these positions. Replacement of A at position 3 by either G or T (and, consequently, T at position 7 by either C

© 1999 Biochemical Society

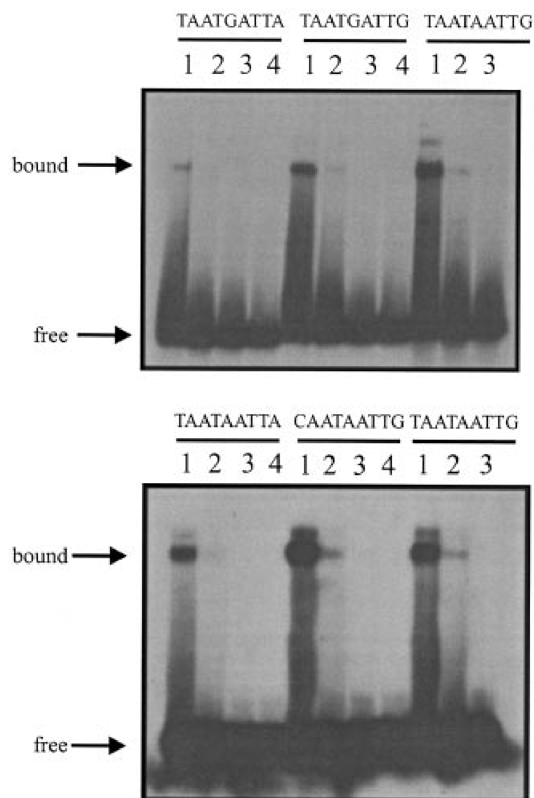


Figure 4 Binding of Hahb-4 to oligonucleotides containing variants of the consensus sequence

Different amounts of Hahb-4 (60, 30 and 15 ng in lanes 1, 2 and 3, respectively) were analysed for binding to 55-mer double-stranded oligonucleotides obtained after cleavage of the corresponding clones with *Eco*RI and *Hin*dIII, and purification through preparative low-melting-point agarose gels (upper panel), or to 24-mer synthetic double-stranded oligonucleotides with either C/G or T/A pairs at positions 1 and 9 of the consensus sequence (lower panel). The central sequences of the oligonucleotides used are shown, all in the same orientations with respect to their arms. In lane 4, the protein was omitted.

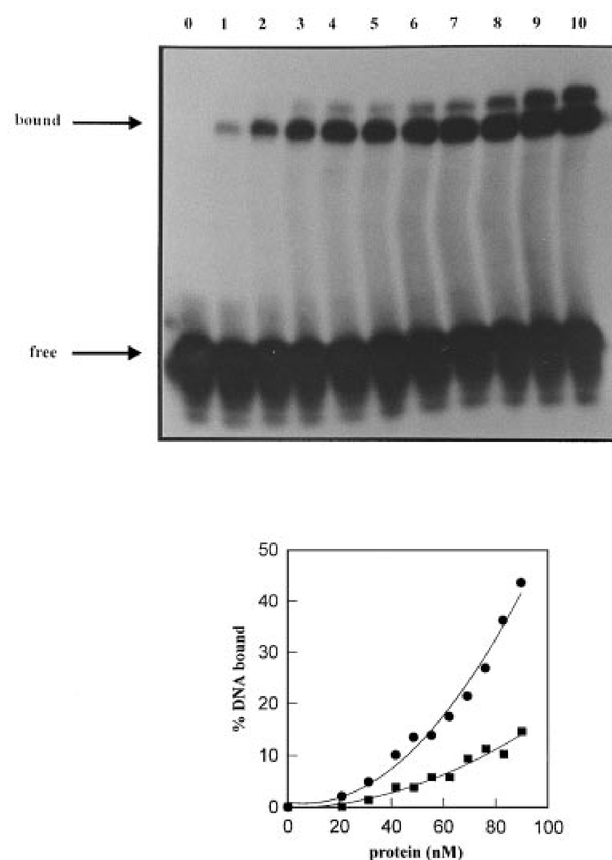


Figure 6 Binding of different amounts of Hahb-4 to DNA

Upper panel: electrophoretic mobility-shift assay using 15–65 ng of Hahb-4 (lanes 1–10, respectively) and a 24-mer double-stranded oligonucleotide containing the central sequence CAATAATTG. Lower panel: quantitative analysis of the percentage of total DNA bound as a function of total protein concentration for this oligonucleotide (●) and for the variant containing the sequence TAATAATTG (■). A 40-kDa monomer was used to calculate the molar concentration of the protein.

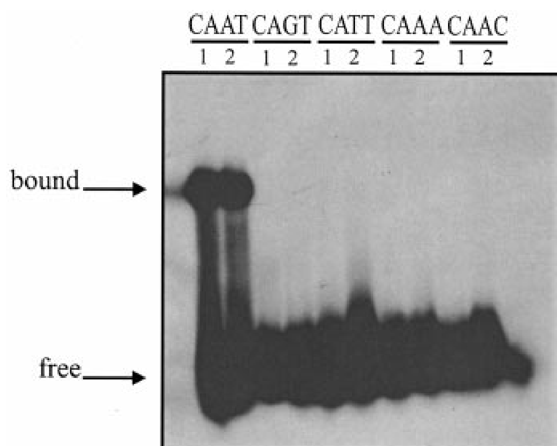


Figure 5 Binding of Hahb-4 to oligonucleotides containing variations at fixed positions

Different amounts of Hahb-4 (50 and 100 ng in lanes 1 and 2, respectively) were analysed for binding to 24-mer synthetic double-stranded oligonucleotides with modifications at either positions 3 and 7 or 4 and 6 of the sequence CAATAATTG. All oligonucleotides contained a dyad-symmetrical sequence and only the first halves (positions 1–4) are indicated.

or A) completely abolished binding (Figure 5). A similar result was obtained when T at position 4 and A at position 6 were replaced by either A/T or C/G, respectively (Figure 5). These results clearly show the importance of these positions for binding, and support the notion that they are contacted by Ile-47 and Asn-51, as described for other homeodomains [2,3].

We conclude then that Hahb-4 shows a preference for the symmetric sequence CAATAATTG, which is similar to that recognized by Athb-1, another class-I Hd-Zip protein. However, single nucleotide changes at positions 1, 5 and 9 did not strongly affect binding affinity. This most probably explains the poor selection obtained at these positions from a random mixture. Similar sequence preferences were observed when either a fusion protein containing the entire C-terminal region, in addition to the Hd-Zip domain, or a protein that had been separated from glutathione S-transferase by cleavage with factor X_a, was used (results not shown).

When different amounts of Hahb-4 were tested for binding to either CAATAATTG or TAATAATTG, a non-hyperbolic response was obtained with both DNAs (Figure 6). A plot of the ratio of bound to free DNA as a function of total protein concentration could be fitted to eqn. (1):

$$PD/D = k \times P_t^2 \quad (1)$$

© 1999 Biochemical Society

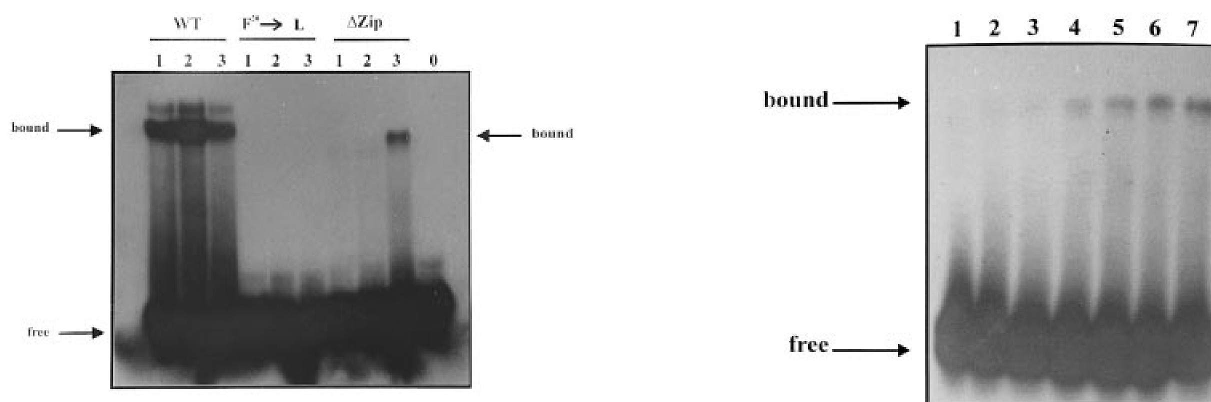
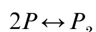


Figure 7 Binding of modified forms of Hahb-4 to different DNA sequences

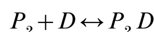
Wild-type (WT) or a Phe-20 → Leu ($F^{20} \rightarrow L$) mutant (60 ng each), and 400 ng of a variant without the leucine-zipper motif (ΔZip) were analysed for binding to oligonucleotides containing the sequences TAATAATTA (lanes 1), CAATAATTG (lanes 2) or TAATAATTG (lanes 3). In lane 0, no protein was included.

where PD and D are the concentrations of bound and free DNA, respectively, P_t is the total protein concentration and k is a constant. Assuming that Hahb-4 interacts with DNA as a dimer, its binding equation can be described by



$$2P \leftrightarrow P_2$$

and



$$P_2 + D \leftrightarrow P_2D$$

with the respective dissociation constants defined as

$$K_1 = P^2/P_2 \quad (2)$$

and

$$K_2 = P_2 \times D/P_2D$$

The global dissociation constant for these two reactions is then

$$K_1 \times K_2 = K_{12} = P^2 \times D/P_2D$$

Since, under the conditions used for the binding assay, $P_t \gg P_2D$, then $P_t = 2 \times P_2 + P$. For $P_t \ll K_1$, $P_t \cong P$ and

$$P_2D/D = P_t^2/K_{12} \quad (3)$$

which is similar to eqn. (1) and shows that, when $P_t \ll K_1$, the ratio of bound to free DNA is proportional to the square of the total protein concentration. If $P_t \gg K_1$, $P_t \cong P_2 \cong P^2/K_1$, and

$$P_2D/D = P_t/K_2 \quad (4)$$

which shows that, when $P_t \gg K_1$, the ratio of bound to free DNA is directly proportional to the protein concentration used in the assay.

We conclude that the type of plot shown in Figure 6 arises from the fact that, under our assay conditions, an important proportion of Hahb-4 is in the monomeric form and that the dissociation constant of dimers (K_1) is well over the protein concentrations used in the assay (20–90 nM of monomers). Fitting eqn. (3) to the experimental results, we calculated values for K_{12} of $1.31(\pm 0.09) \times 10^{-14}$ M^2 for the sequence CAATAATTG and of $5.10(\pm 0.20) \times 10^{-14}$ M^2 for TAATAATTG. Assuming that K_1 is the same, we can conclude that the replacement of C at position 1 for T produces an approx. 4-fold increase in the dissociation constant of the dimer–DNA complex.

Since dimer formation seems to be a pre-requisite for binding, we have tested the effect of removing the leucine-zipper motif. Figure 7 shows that binding to DNA can still be detected, although at higher protein concentrations. Most notably, the homeodomain alone of Hahb-4 shows more affinity for the sequence TAATAATTG than for CAATAATTG, which is preferred by the dimer. This shows that, although low in affinity, the interaction is specific, and that the homeodomain of Hahb-4 spans both moieties of the sequence, since changes at either half produce changes in affinity. A plot of the bound-to-free-DNA ratio at different protein concentrations showed a linear regression ($r^2 = 0.992$), indicating that dimer formation had been abolished (Figure 8). The calculated dissociation constant for this interaction was $5.03(\pm 0.10) \times 10^{-5}$ M.

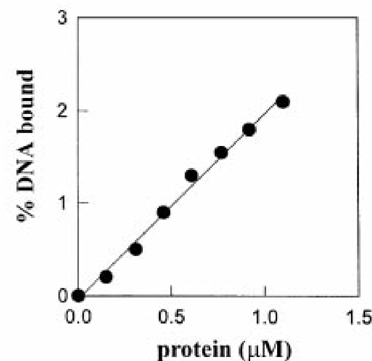


Figure 8 Binding of the Hahb-4 homeodomain to DNA

Different amounts (100–700 ng) of a fusion protein carrying the Hahb-4 homeodomain were used in an electrophoretic mobility-shift assay together with an oligonucleotide containing the sequence TAATAATTG (upper panel). Lower panel, a quantification of the binding assay is shown.

We have also constructed a set of mutants to analyse further the interaction of Hahb-4 with DNA. Replacement of Phe-20 (conserved in almost all homeodomains known so far [1,18]) for Leu abolished binding completely (Figure 7). Since Phe-20, located within helix I, does not directly interact with DNA, but is rather required to maintain the overall conformation of the homeodomain, we conclude that the homeodomain of Hahb-4 should fold into the characteristic three-helix structure observed for other homeodomains. Conversely, the replacement of Arg-5 in the N-terminal arm for Lys had no significant effect on binding (in terms of both affinity and specificity; results not shown).

The availability of a mutant form unable to bind DNA offered the possibility of studying the interaction of monomers through

© 1999 Biochemical Society

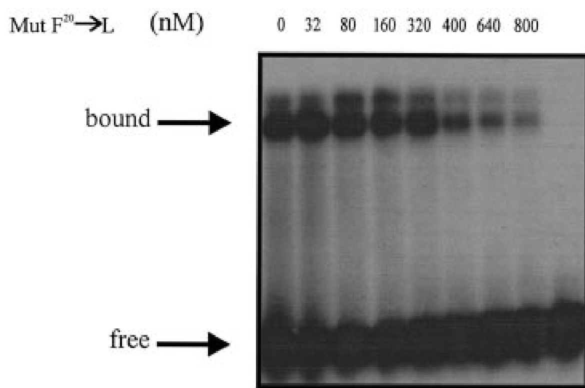


Figure 9 A Phe-20 mutant inhibits binding of wild-type Hahb-4 to DNA

Binding of the Hahb-4 homeodomain leucine-zipper motif (50 ng) to DNA was assayed in the absence or presence of increasing amounts of a Phe-20 → Leu mutant (Mut F^{20} → L, nM). Both proteins were incubated for 15 min at room temperature in the assay mixture before the addition of an oligonucleotide containing the sequence CAATAATTG. The last lane shows an assay with 120 ng of the mutant protein in which the wild-type form was omitted.

the leucine zipper. As shown in Figure 9, the presence of increasing amounts of the Phe-20 mutant inhibited the interaction of native Hahb-4 with DNA. This is most probably explained by the formation of inactive heterodimers between the native and mutant forms (a sort of dominant negative effect). Assuming that dimer formation is not affected by the mutation, the dissociation constant of dimers (K_1, eqn. 2) can be estimated from eqn. (5):

$$P_2D/D = P_t^2/[K_2(K_1 + 2P^* + P^{*2}/K_1)] \quad (5)$$

where P^* represents the concentration of mutant protein and $K_1 \times K_2 = K_{12} = 1.31 \times 10^{-14}$ M^2. The inhibition data fitted well with a value of K_1 within the micromolar range [$(2–4) \times 10^{-6}$ M] and, consequently, the dissociation constant of the dimer–DNA complex fell into the nanomolar range [$(3–6) \times 10^{-9}$ M].

DISCUSSION

In the present study, we analysed the interaction of the Hd-Zip domain of Hahb-4 with DNA. Our studies indicate that, although only 43 % of the amino acids from the Hd-Zip domain are the same as those of Athb-1 (52 % within the homeodomain; 29 % within the leucine zipper), a sequence similar to that bound by Athb-1, another class-I Hd-Zip protein, is preferred by Hahb-4. From these results, it can be postulated that all class-I Hd-Zip proteins described up to now (from which Hahb-4 is the most divergent member) should show similar binding characteristics, at least *in vitro*. This raises the question of the specificity of the interaction of these proteins with target genes *in vivo*. Is this interaction simply governed by the different concentrations of each protein within each cellular type, or are there additional mechanisms (interaction with other proteins, local DNA conformation) that are important in target-gene selection?

Apart from its preferential binding to a dyad-symmetrical sequence of the type CAATAATTG, Hahb-4 showed other interesting characteristics. First, single nucleotide replacements at positions 1, 5 and 9 diminished binding affinity by a factor of 2–4. This, and the fact that the symmetrical sequence could not be selected efficiently from a random mixture, suggests that Hahb-4 is able to bind to several related sequences. Secondly, under the conditions used for the binding assay, a high proportion of the protein remained as monomers, suggesting that the dissociation constant for dimers might be in the micromolar range. In fact, a value of $2–4 \times 10^{-6}$ M could be estimated from the inhibition of DNA binding by a Phe-20 mutant. This situation contrasts with that observed for basic motif leucine-zipper proteins, which show dissociation constants of about 10^{-8} M [24,25]. Since monomers do not seem to bind DNA with considerable affinity, this would mean that Hahb-4 might be very inefficient for binding *in vivo*, unless it is expressed at high concentrations. The significance of this observation is intriguing. It may be that Hahb-4 dimerizes efficiently with another protein, but not with itself, or that *in vivo* homodimer formation is stabilized by additional protein–protein interactions. Alternatively, this may provide a means of regulation, since only when expressed at high levels would DNA-binding entities be produced. In addition, within certain ranges, a sort of co-operative response would be observed, since binding is proportional to the square of protein concentration. It would be interesting to know if other Hd-Zip proteins show a similar behaviour. This mode of regulation is particularly attractive, since Hd-Zip proteins have been proposed as regulators of plant developmental responses to changes in the environment [7], which are rapid and reversible in nature. In the case of Hahb-4, we have detected a strong induction of its mRNA as a consequence of drought stress (G. M. Gago, D. H. Gonzalez and R. L. Chan, unpublished work). Our observations also suggest that mutants unable to bind DNA, such as Phe-20, could be used as dominant negative effectors *in vivo* to study the function of the corresponding protein. From the values of the respective dissociation constants, however, it can be predicted that relatively high concentrations will be required to efficiently compete with DNA binding.

Regarding the mode of binding to DNA, the general assumption made for Athb-1 and -2, that each moiety of the dimer interacts with a half of the dyad-symmetric sequence [16], seems to apply to Hahb-4. An intriguing feature is the asymmetry observed at the central position, which may indicate that one of the monomers (defined by its location with respect to the asymmetric arms) contacts the central base pair more efficiently than the other. This most probably indicates that adjacent sequences influence the interaction, and provide additional specificity.

The presence of a leucine zipper has prompted comparisons of the modes of binding of basic motif leucine-zipper and Hd-Zip proteins to DNA. Residues putatively located at similar positions within the basic region and helix III have been compared, and interesting analogies have been found [16,17]. Some differences, however, are also evident from our studies. First, the fact that mutation of Phe-20 completely abolished binding suggests that a correct alignment of helix III promoted by dimerization is not enough for binding. Therefore, either regions outside helix III are also involved in binding, or helix III becomes disturbed in the mutant at position 20 (helix I). Secondly, the homeodomain becomes folded stably in the absence of the leucine zipper, as suggested by its specific DNA-binding properties. Thus in all respects, the homeodomain behaves as an independent entity, showing specific, although inefficient, DNA binding. It can be proposed that the correct folding of the homeodomain is responsible primarily for its correct alignment with DNA, and that the main function of the leucine zipper would be to join two adjacent homeodomains, so as to increase the number of protein–DNA interactions and, consequently, the binding strength. In this respect, Hd-Zip proteins would be similar to prokaryotic helix-turn-helix proteins, which use dimerization motifs in a similar way.

For several homeodomains, the high affinity for DNA displayed by the monomeric form has been ascribed to the

© 1999 Biochemical Society

existence of specific interactions between amino acids in the N-terminal arm and DNA [1,18,26,27]. Since Hahb-4 contains an Arg at position 5 (a conserved amino acid known to interact with DNA in other homeodomains [1–3,18]), we have tested the effect of its mutation to Lys (found in most Hd-Zip proteins). As shown in the Results section, this substitution has no effect on binding, indicating the lack of a specific interaction of Arg-5 with DNA, perhaps because the overall sequence context is not favourable for binding.

In conclusion, we have observed that Hahb-4 binds preferentially to the sequence CAATAATTG, and to other related sequences with slightly less affinity. Under the conditions used for the assay, a dimer–monomer equilibrium exists, with dimers binding considerably more efficiently than monomers to DNA. Although inefficient, the homeodomain alone is capable of specific DNA binding. The big difference that exists between the dissociation constants for dimer formation and DNA binding suggests that the interaction of this protein with DNA is governed mainly by dimer formation. We propose that these properties may be the basis for efficient regulation of the interaction of Hd-Zip proteins with target genes within the cell.

This work was supported in part by grants from Fundación Antorchas (Argentina). R. L. C. and D. H. G. are members of CONICET; C. M. P. is a Fellow of the same institution.

REFERENCES

1 Gehring, W. J. (1987) Science **236**, 1245–1252
2 Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wüthrich, K. (1994) Cell **78**, 211–223
3 Gehring, W. J., Affolter, M. and Bürglin, T. (1994) Annu. Rev. Biochem. **63**, 487–526
4 Vollbrecht, E., Veit, B., Sinha, N. and Hake, S. (1991) Nature (London) **350**, 241–243
5 Ruberti, I., Sessa, G., Lucchetti, S. and Morelli, G. (1991) EMBO J. **10**, 1787–1791
6 Schena, M. and Davis, R. W. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 3894–3898
7 Chan, R. L., Gago, G. M., Palena, C. M. and Gonzalez, D. H. (1998) Biochim. Biophys. Acta **1442**, 1–19
8 Bharathan, G., Janssen, B. J., Kellog, E. A. and Sinha, N. (1997) Proc. Natl. Acad. Sci. U.S.A. **94**, 13749–13753
9 Bürglin, T. R. (1997) Nucleic Acids Res. **25**, 4173–4180
10 Mattsson, J., Söderman, E., Svenson, M., Borkird, C. and Engström, P. (1992) Plant Mol. Biol. **18**, 1019–1022
11 Chan, R. L. and Gonzalez, D. H. (1994) Plant Physiol. **106**, 1687–1688
12 Schena, M. and Davis, R. W. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 8393–8397
13 Carabelli, M., Sessa, G., Baima, S., Morelli, G. and Ruberti, I. (1993) Plant J. **4**, 469–479
14 Gonzalez, D. H., Valle, E. M., Gago, G. M. and Chan, R. L.(1997) Biochim. Biophys. Acta **1351**, 137–149
15 Meijer, A. H., Scarpella, E., van Dijk, E. L., Qin, L., Taal, A. J. C., Rueb, S., Harrington, S.E., McCouch, S. R., Schilperoort, R. A. and Hoge, J. H. C. (1997) Plant J. **11**, 263–276
16 Sessa, G., Morelli, G. and Ruberti, I. (1993) EMBO J. **12**, 3507–3517
17 Sessa, G., Morelli, G. and Ruberti, I. (1997) J. Mol. Biol. **274**, 303–309
18 Laughon, A. (1991) Biochemistry **30**, 11357–11367
19 Smith, D. B. and Johnson, K. S. (1988) Gene **67**, 31–40
20 Palena, C. M., Gonzalez, D. H., Guelman, S. and Chan, R. L. (1998) Protein Expression Purif. **13**, 97–103
21 Oliphant, A. R., Brandl, C. J. and Struhl, K. (1989) Mol. Cell Biol. **9**, 2944–2949
22 Blackwell, T. K. and Weintraub, H. (1990) Science **250**, 1104–1110
23 Sedmak, J. and Grossberg, S. (1977) Anal. Biochem. **79**, 544–552
24 Patel, L. R., Curran, T. and Kerppola, T. K. (1994) Proc. Nat. Acad. Sci. U.S.A. **91**, 7360–7364
25 Krylov, D., Olive, M. and Vinson, C. (1995) EMBO J. **14**, 5329–5337
26 Shang, Z., Isaac, V. E., Li, H., Patel, L., Catron, K. M., Curran, T., Montelione, G. T. and Abate, C. (1994) Proc. Nat. Acad. Sci. U.S.A. **91**, 8373–8377
27 Damante, G., Pellizzari, L., Esposito, G., Fogolari, F., Viglino, P., Fabbro, D., Tell, G., Formisano, S. and Di Lauro, R. (1996) EMBO J. **15**, 4992–5000

Received 24 November 1998/16 March 1999; accepted 15 April 1999

© 1999 Biochemical Society