

# Cooperative dimerization of Paired class homeo domains on DNA

David Wilson, Guojun Sheng, Thomas Lecuit, Nathalie Dostatni, and Claude Desplan

Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021-6399 USA

**Homeo domain-containing proteins mediate many transcriptional processes in eukaryotes. Because nearly all animal homeo proteins are believed to bind to short, highly related DNA sequences, the basis for their high specificity of action is not understood. We show that cooperative dimerization on palindromic DNA sequences can provide increased specificity to one of the three major classes of homeo domains, the Paired/Pax class. The 60-amino-acid homeo domains from this class contain sufficient information to bind cooperatively as homo- and heterodimers to palindromic DNA sequences; that is, the binding of one homeo domain molecule can increase the affinity of a second molecule by up to 300-fold. Different members of the Paired (Prd) class of homeo domains prefer different spacings between half-sites, as determined by the ninth amino acid residue of the recognition helix. In addition, this residue determines the identity of the base pairs at the center of the palindromic sites, as well as the magnitude of the cooperative interaction. The cooperative dimerization of homeo domains in the Prd class distinguishes them from other classes, whereas binding-site configuration and sequence specificity allow for distinctions within this class.**

[*Key Words:* Cooperativity; dimerization; DNA binding; homeo domain; recognition helix; transcription]

Received July 9, 1993; revised version accepted August 26, 1993.

The homeo domain (HD) is found in multiple gene products within an organism and across eukaryotic phylogeny, where it plays an important role in developmental programs (Scott et al. 1989; McGinnis and Krumlauf 1992). HD swap experiments have shown that the functional targeting differences between homeo proteins are largely determined by the HDs themselves (Kuziora and McGinnis 1989; Gibson et al. 1990; Hayashi and Scott 1990; Mann and Hogness 1990). The HD can recognize specific DNA sequences, but it is not known whether different functions between homeo proteins result from differences in DNA-binding specificity (Hanes and Brent 1989, 1991; Treisman et al. 1989, 1992; Dessain et al. 1992; Ekker et al. 1992; Schier and Gehring 1992, 1993).

The HD recognizes DNA through a region of homology to the helix–turn–helix motif present in a group of prokaryotic transcription factors (Laughon and Scott 1984; Qian et al. 1989). In both the prokaryotic and eukaryotic cases, the second helix of this motif recognizes bases in the major groove of DNA (Pabo and Sauer 1992). However, genetic (Hanes and Brent 1989, 1991; Treisman et al. 1989; Percival-Smith et al. 1990) and structural (Kissinger et al. 1990; Otting et al. 1990; Wolberger et al. 1991) studies have revealed that whereas the prokaryotic recognition helix utilizes its amino-terminal six residues to interact with DNA (Pabo and Sauer 1992), the HD utilizes a different geometry with respect to the DNA, and specificity relies on the ninth residue. Structural and biochemical comparisons between different HDs have converged on the following model for interac-

tions between higher eukaryotic HDs and their binding sites: Conserved residues of the HD amino terminus and of the recognition helix contact a **TAAT** core in both the major and minor grooves, whereas different ninth residues of the recognition helix interact with and specify the 2 bp immediately 3' to this core. This model has been confirmed with several examples (Treisman et al. 1992) and supported by *in vivo* manipulations of the homeo protein Fushi tarazu (Ftz) (Furukubo-Tokunaga et al. 1992; Schier and Gehring 1992).

Although most DNA-binding domains bind either as dimers (prokaryotic helix–turn–helix, nuclear receptor family, bZIP, helix–loop–helix, etc.) or as higher order complexes (heat shock factor, Lac repressor), or possess multiple recognition domains (C<sub>2</sub>H<sub>2</sub> zinc finger proteins) (Freemont et al. 1991; Pabo and Sauer 1992), the 60-amino-acid HD binds identified target sites as a monomer (Kissinger et al. 1990; Otting et al. 1990; Affolter et al. 1991; Wolberger et al. 1991). As expected by this limitation, HD recognition sequences consist of only 5–6 bp (Treisman et al. 1992), a paradoxical attribute for a domain that has evolved so many distinct functional specificities. Furthermore, several studies have suggested that highly diverged HDs have indistinguishable DNA-binding preferences (Desplan et al. 1988; Hoey and Levine 1988). Thus, the apparent lack of specificity in DNA binding suggests that other mechanisms are required to achieve the individual functions of homeo proteins (Hayashi and Scott 1990). Recently described examples of such mechanisms include homeo protein oli-

gomerization, mediated by dimerization domains outside of the HD (Dranginis 1990; Ingraham et al. 1990; Rey-Campos et al. 1991), as well as protein–protein interactions of homeo proteins with other factors (Keleher et al. 1988; Grueneberg et al. 1992; Lai et al. 1992; Pomerantz et al. 1992; Smith and Johnson 1992), and association of the HD with other DNA-binding domains (Ingraham et al. 1990; Chalepakis et al. 1991; Fortini et al. 1991; Ruberti et al. 1991; Treisman et al. 1991; Verrijzer et al. 1992). In contrast to these cases in which protein regions other than the HD itself are required, we describe here a new mechanism by which the 60-amino-acid HD can itself confer highly specific interactions with DNA.

We chose to investigate the specificity determinants of one of the three largest classes of homeo proteins, the Paired (Prd) class. Within this class are members that also possess a second DNA-binding motif, the Prd domain [e.g., *Drosophila* Prd, Pax HD genes (Bopp et al. 1986; Gruss and Walther 1992)], and other members that do not [e.g., *Drosophila* Orthodenticle (Finkelstein et al. 1990), *Xenopus* Goosecoid (Gsc) and Mix-1 (Rosa 1989; Blumberg et al. 1991), and human Phox-1 (Grueneberg et al. 1992)]. Disruption of Prd class homeo proteins results in developmental abnormalities in a wide range of organisms including the nematode (Miller et al. 1992), *Drosophila* (Baumgartner et al. 1987), mouse (Epstein et al. 1991; Hill et al. 1991), and human (Gruss and Walther 1992). Because little is known about the DNA binding specificity of the Prd class of homeo proteins and no biological target sites have been identified for any of its members, we chose to employ the SELEX method (Thiesen and Bach 1990; Tuerk and Gold 1990; Szostak 1992), in which optimal DNA ligands are isolated from random oligonucleotide libraries. This approach revealed that the Prd class HDs bind optimally to palindromic DNA sequences composed of two TAAT half-sites. The 60-amino-acid HD contains sufficient information to cooperatively dimerize on these sites, thus distinguishing the HDs of the Prd class from those of other classes, which bind as monomers. We also demonstrate that the spacing between half-sites and, hence, the structure (conformation) of the DNA-binding complex, varies between members of the Prd family. This conformational specificity is conferred by the ninth residue of the HD recognition helix: When the position is occupied by a serine (as in *Drosophila* Prd), the structure of the cooperative dimer on DNA is more compact than when it is a glutamine (as in Mix-1) or a lysine (as in Gsc). In addition, this residue also determines the identity of the base pairs at the center of the palindromic sites, as well as the magnitude of the cooperative interaction. Therefore, HD cooperative dimerization, binding-site configuration, and binding-site sequence specificity allow for distinctions between homeo proteins within the Prd class.

## Results

### *The Prd HD binds as a cooperative dimer to palindromic DNA sequences*

We have designed a simple and general version of the

SELEX procedure (Thiesen and Bach 1990; Tuerk and Gold 1990; Szostak 1992) to select optimal ligands for DNA-binding proteins. A double-stranded DNA oligonucleotide containing 18 random positions in the center, flanked by constant regions for PCR amplification, is incubated with a fusion protein between glutathione-S transferase (GST) and one of various HDs. DNA associated with such beads is amplified by PCR, and the resulting DNA is then entered into further generations of selection and amplification. Individual oligonucleotides selected by the HD are then cloned and sequenced, and sequences are aligned to generate a consensus. We initially used this technique on two HDs from *Drosophila* proteins (but not from the Prd class) that have been shown to interact in vivo with specific promoter sites. Both the Ftz and the Bicoid (Bcd) HDs selected sites nearly identical to those functionally recognized in vivo [TAATCC for Bcd and TAATGG for Ftz (Driever and Nüsslein-Volhard 1989; Shier and Gehring 1992)]. As predicted by the current model (see introductory section) of HD–DNA interaction and specificity, they selected identical core TAAT elements but preferred different base pairs immediately downstream of this, according to the differences between the two HDs at recognition helix position 9 (lysine for Bcd, glutamine for Ftz).

We then performed the binding-site selection on the HD from the Prd protein, for which no biological target sites are known. The sequences selected by the Prd HD also all contain the core HD binding motif TAAT. They fall into a major and a minor group (Fig. 1A). Figure 1B represents the consensus from the more common type of sequence (the minor group will be discussed below). Unexpectedly, the consensus is a palindrome composed of two inverted TAAT sites separated by 2 bp (therefore called P2). The symmetric nature of the consensus suggested that two Prd HD molecules bind to the sequence. To determine the stoichiometric nature of the complex between the Prd HD and the P2 sequence, and to eliminate any potential involvement of GST, we overproduced a 65-amino-acid HD peptide in *Escherichia coli* and tested the extract for DNA binding in a mobility shift assay. Figure 2A shows that the Prd HD peptide binds the P2 site (TAATTGATTA) to produce a major complex (D) that migrates slower than a less abundant complex (M). On a probe containing only one half-site (P1/2, TAATTG), the HD shifts the probe only to a position corresponding to the lower complex on P2 (M). This suggests that the Prd HD binds DNA predominantly as a dimer (D) on the P2 probe, and, with lower affinity, as a monomer (M) when only one half-site is supplied. The dimeric nature of the major species bound to P2 was confirmed by mixing a higher and lower molecular weight Prd HD construct (Hope and Struhl 1987), and observing an intermediate band, analogous to the experiment shown in Figure 2B (data not shown). Thus, it appears that the Prd HD can bind to a single TAAT half-site and that one monomer binding event on P2 is capable of facilitating a second one. We refer to this two-step process as cooperative dimerization. Qualitatively, the cooperativity can be seen from the fact that at con-

Wilson et al.

A

**major consensus:**

TGATGTGCTG**TAAT**TGATTACATGAT  
 TTCATGATA**TAAT**CGATTATTGATGTGGTGAT  
 GAGTGGGGCG**TAAT**TGATTAACTGAT  
 TGATTGCGGTGCT**TAAT**TGATTATGATATCG  
 TGATTCGGCC**TAAT**TGATTAGTTGATATCG  
 TTCATGATA**TAAT**TGATTACGCTTCCCTGAT  
 ATCAAGCCGTA**CTTAAT**TGATTATCATGAA  
 TGATGCT**TAAT**CGATTACGGGCTGAT  
 TTCATGATA**TAAT**CGATTAGCACGGTGTGAT  
 TGATGGCG**TAAT**CGATTAAAGTCATGAT  
 ATCAGGCAGAGGT**TAAT**TGATTATCATGAA  
 TTCATGATA**TAAT**TGATTACCTTGCTGTTGAT  
 TTCATGATA**TAAT**CGATTATGGGGCGCTTGAT  
 TGATTAGCT**TAAT**TGATTAGCGGTGAT  
 TTCATGATA**TAAT**TGATTAGCGTCAGGCTGAT  
 ATCACGACAAAATCCG**TAAT**CGATCATGAA  
 ATCAACATCCCG**TAAT**TGATTATCATGAA  
 TGATGCATTTTGACTTGATA**TAAT**TGATTTCGAAGT  
 TTCATGATA**TAAT**CGATTAGTACGCTTTGAT

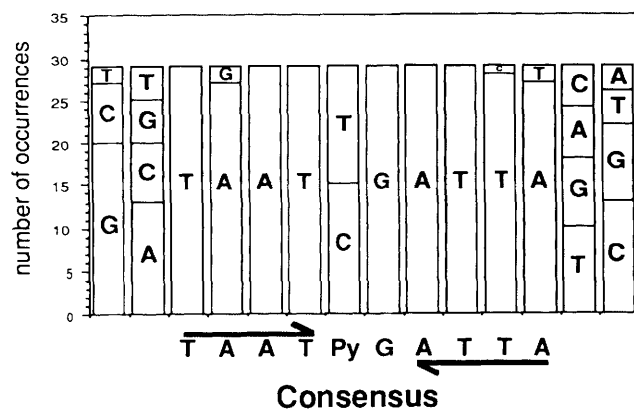
TTCATGATCGCT**TAAT**CGATTAAAGCGGTGAT  
 TTCATGATA**TAAT**CGATTAGTATACGGGTGAT  
 TGATGTGCG**TAAT**CGATTAACTTGAT  
 ATCAACGGCGCT**TAAT**TGATTATCATGAA  
 TTCATGATA**TAAT**CGATTACGTTGCCTATGAT  
 TGATATT**TAAT**CGATTAGCTGTTGAT  
 ATCAGGGCC**TAAT**TGATTACAAATCA  
 TTCAT**TAAT**CGATTATAGTCCGTTGGTAT  
 TTCAT**TAAT**CGATTAACTTCGCATACCTGAT  
 TTCATGATA**TAAT**CGATTAGGCTTAGGGTGAT

**minor consensus:**

TTCAT**TAAT**CGATTAGGGGTGTAATGTGAT  
 TGATACGTGCGT**TAAT**CGAATTAATGATATCG  
 TTCAT**TAAT**CGAATTAGCGCGATTAGTTGAT  
 ATCAATTAGCGGCAC**TAAT**TAGATCATGAA  
 TGATGCGTT**TAAT**CCGGTTACCTGATATCG  
 TTCATGATA**TAAT**CGATTAGTCTCCCTGAT

**Figure 1.** DNA-binding site consensus for the Prd HD. (A) Alignment of random oligonucleotides selected by the Prd HD. Of the 37 sequenced, 29 aligned according to the major consensus, 6 to the minor consensus and 2 contained **TAAT** motifs that could not be assigned unambiguously to either group. For each sequence, the **TAAT** core binding site is in bold-face type. (B) "Consensogram" of the major consensus selected by the Prd HD. The height of each stacked bar represents the number of occurrences of each base at each position along the consensus. The consensus is a palindrome of two **TAAT** motifs separated by 2 bp and is therefore called P2.

B



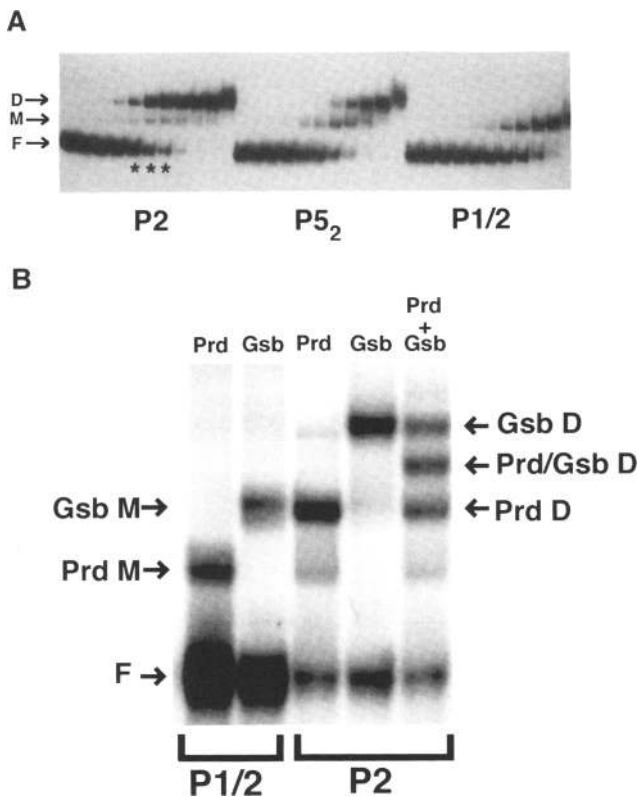
centrations of HD where relatively little of the probe is shifted, more dimerically (D) than monomerically (M) bound probe is present. This is in contrast to the way in which the Prd HD binds to the P5<sub>2</sub> probe, in which 5 rather than 2 bp separate the **TAAT** half-sites (Fig. 2A). We reasoned that the 5-bp separation should disable the cooperativity, because this configuration was not selected from the random DNA pool. P5<sub>2</sub> can also be bound by a single or by two Prd HD molecules, depending on HD concentration. Unlike the case of Prd HD binding to P2, however, the first site of P5<sub>2</sub> becomes nearly saturated before the second one begins to fill (Fig. 2A). Quantitatively, the strength of the cooperative interaction on P2 can be expressed as the cooperativity constant,  $\tau$ , defined as the extent to which the first HD bound to one half-site increases the affinity of a second HD for the adjacent half-site (see Materials and methods). The cooperativity constant can be calculated by comparing the relative amount of HD required to produce an equal amount of unbound and dimerically bound probe on P2 versus P5<sub>2</sub> (using equation 4, Materials and methods). By use of the data in Figure 2A,  $\tau$  is calculated to be 44. That is, the first DNA-binding event on P2 aids

the second binding 44-fold. Another method for determining the cooperativity is presented below and yields similar results. Thus, unlike HDs characterized previously, which have been shown to be capable of recognizing DNA only as monomers, the Prd HD prefers to bind a palindromic DNA sequence as a cooperative dimer.

*Homo- and heterodimerization of Prd class HDs*

We wished to determine whether other Prd class HDs could dimerize on the P2 site and whether cooperative complexes between different HDs could form. Therefore, we mixed the Prd HD peptide with a higher molecular weight peptide containing another Prd class HD (Hope and Struhl 1987). As shown in Figure 2B, the Gooseberry—distal(Gsb) (Baumgartner et al. 1987) HD peptide exhibits behavior similar to Prd on P2 versus P1/2 (or versus P5<sub>2</sub>, data not shown). When the Gsb HD is mixed with the Prd HD, three types of dimeric complexes form on P2, representing the two types of homodimers (Prd-Prd and Gsb-Gsb) and the heterodimer of intermediate mobility (Prd-Gsb). The cooperative nature of the heterodimer on P2 is shown by its ability to compete with





**Figure 2.** Cooperative homo- and heterodimerization of Prd class HDs on DNA. (A) Mobility shift of Prd HD peptide on P2 (TAATTGATTA), P5<sub>2</sub> (TAATTGATGATTA), and P1/2 (TAATTG). For each probe, the first lane contains no HD, and the remaining lanes show the result of increasing amounts (3× difference between each lane) of Prd HD extract from left to right. Equivalent amounts of protein are used for each probe. The concentration of total HD in the lane with the lowest amount of HD is estimated at 0.2 nM. The free probe (F) is shifted to the dimer position (D) on P2, compared with the monomer position (M) on P1/2. Calculations based on the comparison between P2 and P5<sub>2</sub> show that the first binding event on P2 facilitates the second one by 44 times. (B) In a mobility shift experiment, both the Prd HD peptide and the slightly larger Gsb HD peptide shift the P2 probe to a higher position (Prd D and Gsb D, dimers) than they shift the half-site P1/2 (Prd M and Gsb M, monomers). When the two peptides are mixed together with the P2 probe, an intermediate band appears (Prd/Gsb D), representing the heterodimer. The heterodimer is cooperative, as shown by its ability to compete with the Prd/Prd homodimer and by its increased abundance on the P2 probe compared with the P5<sub>2</sub> probe (not shown). The asterisks are referred to in the text.

the two cooperative homodimers, and by the fact that it is present to a much lesser degree on P5<sub>2</sub> at the same concentration of HDs (data not shown). We have also observed cooperative dimerization on P2 with a recently cloned highly divergent *Drosophila* member of the Prd HD class, Lune (S. Jun, B. Kalionis, and C. Desplan, unpubl.). Because this HD is one of the most divergent in the Prd class (only 58% homologous to Prd), it is likely

that the dimerization properties of the Prd HD are conserved within the class.

#### *Cooperative binding results in synergistic activation of transcription*

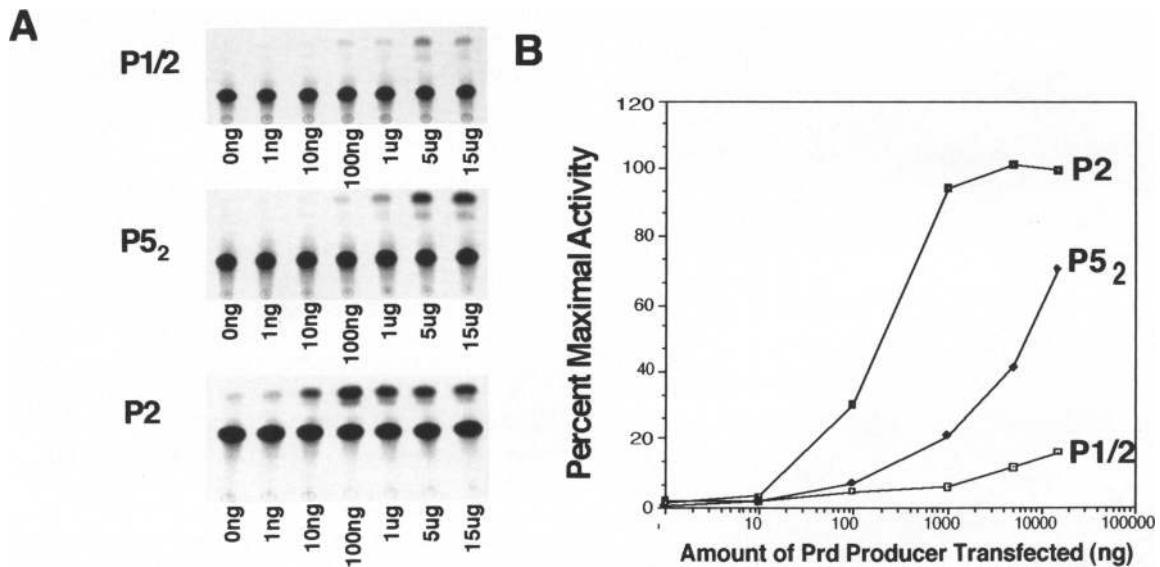
We used cultured Schneider insect cells to test whether the cooperative dimerization observed *in vitro* could lead to synergistic *trans*-activation. We compared the response of promoters containing a single P2, P5<sub>2</sub>, or P1/2 site to Prd full-length protein. Increasing amounts of Prd producer plasmid were cotransfected with a constant amount of a reporter plasmid containing target sites driving CAT. Figure 3A shows that at intermediate levels of protein, P2 responds much more effectively to Prd than does P5<sub>2</sub> or P1/2. Quantitative analysis (Fig. 3B) shows that the amount of Prd producer required to bring about half maximal activation of the promoter with P2 is 40-fold less than with P5<sub>2</sub>. Because both palindromic sites (P2 and P5<sub>2</sub>) support similar levels of activation at high protein concentrations, when nearly all sites are presumably occupied, it is likely that the increased responsiveness of P2 over P5<sub>2</sub> at low protein concentration is attributable to cooperative binding, rather than to different activation potentials of the P2 versus P5<sub>2</sub> bound promoters. Since higher activation levels can be achieved with the addition of a second HD palindromic site (data not shown), saturation of transcriptional activation on P2 results from full site occupancy rather than from transcriptional saturation. These results suggest that in the context of the full-length Prd protein, the HD can cooperatively bind the P2 site *in vivo* and lead to synergistic *trans*-activation.

#### *Residue 9 of the recognition helix alters cooperative binding by the HD*

The HDs that cooperatively dimerize on the P2 site, Prd, Gsb, and Lune, all possess a serine at the ninth position of their recognition helices. This position in other HDs has been shown previously to interact with the base pairs immediately 3' to the TAAT core sequence (Hanes and Brent 1989, 1991; Kissinger et al. 1990; Otting et al. 1990; Percival-Smith et al. 1990; Treisman et al. 1992). These base pairs correspond to the center of symmetry in the P2 palindrome. We therefore investigated how the identity of this amino acid would influence the cooperative dimerization of Prd class HDs. Because other members of the Prd class possess a glutamine or a lysine at this position, we performed the random-site selection assay with the Prd HD in which the serine nine (S9) had been mutated to glutamine (Q9) or lysine (K9).

All of the sequences selected by Prd HD Q9 display a motif allowing alignment (Fig. 4A). As with the wild-type Prd HD, this consensus contains two inverted TAAT sequences, each followed by a pyrimidine. In the case of the Q9 mutant, however, 3 bp instead of two separate the inverted TAATs. Figure 4B shows the 65-residue Prd HD Q9 peptide cooperatively binding to the P3 sequence (3 bp separating the inverted TAATs) but

Wilson et al.



**Figure 3.** Cooperative binding of the Prd HD on the P2 site strengthens the *trans*-activation potential of Prd in *Drosophila* cell culture. (A) CAT assay of Prd activity on P2, P5<sub>2</sub>, and P1/2. For each binding site, increasing amounts of Prd producer plasmid and equal amounts of reporter were added, as indicated. (B) Quantification of the experiment shown in A. CAT activity is expressed in arbitrary units after normalization of transfection efficiency with a vector expressing  $\beta$ -galactosidase. The maximal activity on the P2 site is designated as 100. The P2 site requires 40 $\times$  less producer to reach half-maximal activity than does the P5<sub>2</sub> site.

not to a sequence in which 5 bp separate the two inverted TAATs (P5<sub>3</sub>). Binding of the HD to a single half-site (P1/2) marks the position of the singly bound probe. We measured the cooperativity constant,  $\tau$ , by comparing binding of the Prd HD Q9 peptide to the P3 versus P5<sub>3</sub> probes (by equation 4). Using the data from Figure 4B,  $\tau$  is calculated to be 240. Thus, the cooperativity of the Q9 mutant on P3 is much stronger than that of the wild-type (S9) Prd HD on P2. Therefore, the S  $\rightarrow$  Q mutation alters both the relative position of the two HD monomers in the optimal binding complex and the strength of the cooperativity.

The HD dimeric complex formed by Prd HD Q9 on P3 presents an interesting feature: In a mobility shift, it migrates slightly faster than the noncooperative dimer on P5<sub>3</sub> (arrows, Fig. 4B). Mixing two Prd HD Q9 peptides of different molecular weights showed that this P3 complex is a dimer (data not shown). The unusual migration of the cooperative dimer may therefore represent structural changes in the HD–DNA complex. Such changes are not present in the noncooperative P5<sub>3</sub> complex or in the cooperative complex of Prd HD on P2 (see Fig. 2A).

As shown in Figure 4A, the Prd HD Q9 mutant exclusively selected P3 sites, whereas the wild-type Prd HD predominantly selected P2 sites, but also, at a lower frequency, sites related to P3 (Fig. 1A, minor consensus). We tested whether the wild-type Prd HD would bind cooperatively to the palindromic site selected by the Prd HD Q9 (P3), and vice versa, whether Prd HD Q9 would bind cooperatively to the P2 sequence. As predicted by the selection, the Prd HD also binds cooperatively to the P3 sequence, although with an affinity of about one-half of that on P2 (data not shown). The other Prd class mem-

bers Gsb and Lune, which both have a serine at position 9 of the recognition helix, behave similarly to Prd HD (data not shown). In contrast, the Prd HD Q9 does not bind cooperatively to the P2 sequence (Fig. 5A; data not shown), consistent with the selection. The Prd HD Q9 therefore has a more stringent binding specificity than Prd HD. A heterodimer between Prd HD and Prd HD Q9 has an intermediate specificity; that is, it binds cooperatively to both P3 and, to a lesser extent, to P2 (data not shown).

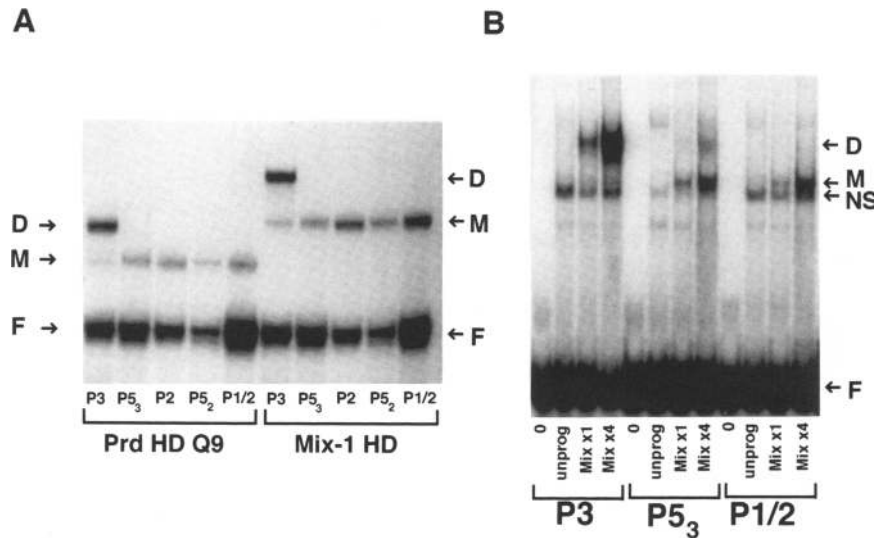
Random-site selections also reveal differences in TAAT half-site binding specificity between the Prd HD and Prd HD Q9. Figure 1 shows that the Prd HD can accommodate TG, CA, or CG following the TAAT core. The absence of the sequence TAATTAATTA from the Prd HD selection suggests that the half-site TAATTA is unacceptable for Prd HD binding. By comparison, this half-site appears to be acceptable for Prd HD Q9 binding, as inferred from Figure 4A and results (not shown) of selections for optimal single TAAT half-sites for Prd HD Q9. Comparison of Figures 2A and 4B also shows that the Prd HD Q9 binds with about a fivefold higher affinity than Prd HD to the half-site TAATTG under the conditions used.

Many HDs in the Prd class possess a Q at position 9 of their recognition helices, such as *Xenopus* Mix-1 (Rosa 1989). We tested whether Mix-1 would bind cooperatively to the P3 sequence, as does the Prd HD Q9 mutant. Mix-1 HD behaves indistinguishably from Prd HD Q9 on the five probes P3, P5<sub>3</sub>, P2, P5<sub>2</sub>, and P1/2, binding cooperatively only to the P3 sequence (Fig. 5A). The cooperativity constant,  $\tau$ , for Mix-1 HD binding to P3 is estimated at 300 (data not shown). Figure 5B shows that





**Figure 5.** Binding specificity of Prd HD Q9 and Mix-1 HD and full-length protein. (A) Mobility shift of Prd HD Q9 and Mix-1 HD on P3 (TAATTGAATTA), P5<sub>3</sub> (TAATTGAGAATTA), P2 (TAATTGATTA), P5<sub>2</sub> (TAATTGATGATTA), and P1/2 (TAATTG). (F) The position of the free probes. At this concentration, Prd HD Q9 binds as a dimer (D) to the palindromic sequence that it selected, P3, and as a monomer (M) to the remaining probes. Mix-1 HD, which naturally possesses a Q at the ninth position of the recognition helix, binds indistinguishably from the Prd HD Q9 mutant. (B) DNA binding of full-length Mix-1 protein produced by in vitro transcription/translation from a reticulocyte extract (RE). (F) Free probes; (0) no RE addition. Addition of 1  $\mu$ l of RE programmed with the RNA from cloning vector without insert (unprog) shows a non-specific shift (NS) on all probes. Addition of 0.25  $\mu$ l (Mix  $\times$ 1) or 1  $\mu$ l (Mix  $\times$ 4) of RE programmed with Mix-1 RNA shows a specific shift on all three probes; the shift is at a higher position (D, dimer) and more abundant on P3, however, than on P5<sub>3</sub> or P1/2 (M, monomer). A faint dimeric complex also appears on P5<sub>3</sub> at the higher Mix-1 concentration.



protein, we then performed a mobility shift experiment (Fig. 7A) at very high probe concentration (0.5  $\mu$ M) to measure precisely the active HD concentration by titration of an accurately determined amount of DNA. The binding curve shown in Figure 7B demonstrates that the amount rather than the concentration/affinity of HD is limiting for probe binding, as expected at this DNA concentration. Assuming a HD/DNA ratio of 2 : 1 for the upper band (D), we found the concentration of active HD to be similar to that estimated by a Bradford assay. We then performed a similar mobility shift with the purified Prd HD Q9 at a low probe concentration (0.2 nM, Fig. 7C) and known HD concentration. In this case, the dose response curve (Fig. 7D), as well as calculations of free versus bound HD concentration (not shown), indicate proper conditions for affinity measurement. In the lanes (marked by asterisks) of Figure 7C in which free probe, monomerically bound probe, and dimerically bound probe can be measured accurately, the two dissociation constants,  $K_{d1}$  (for the unbound  $\rightarrow$  singly bound P3 probe reaction) and  $K_{d2}$  (for the singly bound P3 probe  $\rightarrow$  doubly bound P3 probe reaction) can be calculated (see Materials and methods, equations 1 and 2). Accordingly,  $K_{d1}$  is  $9.7 \pm 2.0$  nM and  $K_{d2}$  is  $40 \pm 0.8$   $\mu$ M. In each lane,  $\tau$  is equal to  $K_{d2}/K_{d1}$  (by definition) and is calculated to be  $246 \pm 32$ . This is similar to the value ( $\tau = 240$ ) determined with the crude cell lysates by comparing binding to P3 versus P5<sub>3</sub>, above (using equation 4). It is also in reasonable agreement with the value of  $\tau$  calculated by considering the relative band intensities for lanes marked by asterisks in Figure 7A; here  $\tau = 186 \pm 10$  (according to equation 3). Therefore, the relative tendency of dimeric as opposed to monomeric binding of the Prd HD Q9 to the P3 probe is unchanged over at least a 3000-fold range in HD concentration. This

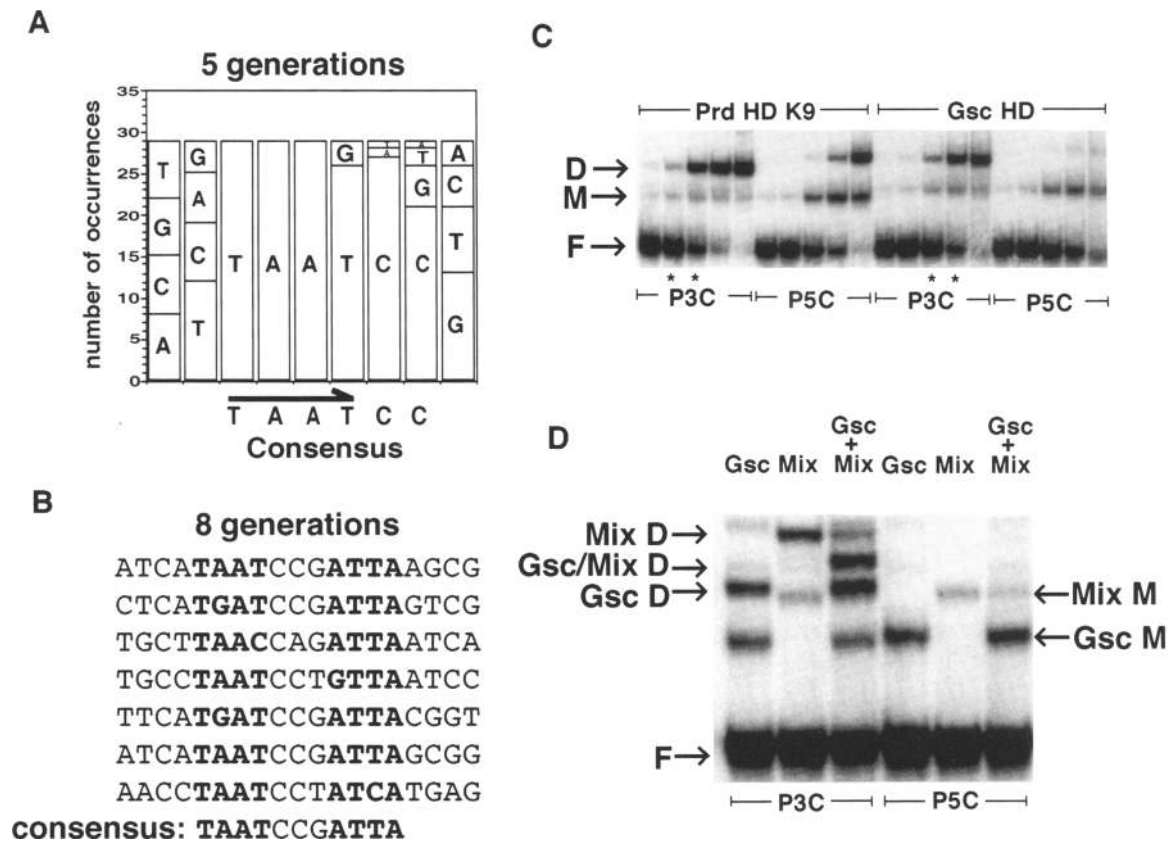
rules out the possibility that the dimerization of the HD on DNA is mediated through a mechanism involving dimer formation in solution. Rather, it shows that the magnitude of the interaction between the two HDs on the palindromic site is truly cooperative, that is, independent of HD concentration.

The cooperativity constants for Prd HD on P2 and Prd HD K9 on P3C were calculated above, on the basis of comparison with noncooperative probes, using equation 4 (Materials and methods). The constant can also be measured by comparing the relative band intensities (cf. F with M with D in Figs. 2A and 6C) in each lane of the mobility shift, and using equation 3. This has the advantage that it is internally controlled; that is, it does not rely on comparisons between different probes or different lanes in a mobility shift. It has the disadvantage that rather faint bands must be quantified. Calculations based on relative band intensities in the lanes marked by asterisks in Figure 2A (where all three bands are clearly visible), show that  $\tau = 65 \pm 26$  for the Prd HD cooperativity on P2. This is in reasonable agreement with the value obtained above ( $\tau = 44$ ), using equation 4 and the comparison between P2 and P5<sub>2</sub>. The same calculation yields  $\tau = 28 \pm 4$  for Prd HD K9 cooperativity on P3C, and  $22 \pm 4$  for Gsc HD cooperativity on P3C, similar to the value obtained above ( $\tau = 23$ ) for Prd HD K9 binding to P3C (by comparison with P5C, using equation 4).

## Discussion

### *Prd family HDs cooperatively bind DNA as homo- and heterodimers*

Prd belongs to one of the largest and most intensely investigated classes of homeo proteins (Baumgartner et al.



**Figure 6.** Prd HD K9 and Gsc HD display similar DNA-binding specificities and cooperativity. (A) Consensogram for Prd HD K9 after five rounds of selection, showing a monomeric HD-binding site. (B) Alignment of 7 of the 17 oligonucleotides sequenced after eight generations of selection by Prd HD K9. The consensus is TAATCCGATTA. (C) Cooperative dimerization of Prd HD K9 and Gsc HD, which both possess a lysine at recognition helix position 9 (K9), on P3C vs. P5C. For each probe, the first lane contains no HD, and the remaining lanes show the result of increasing amounts (3× difference between each lane) of HD extract from left to right. Equivalent amounts of each protein are used for both probes. (F) The free probe; (M) the monomeric HD-DNA complex; (D) the HD dimer on DNA. Comparison between P3C and P5C show that the first binding event on P3C facilitates the second one by 20–30 times for both Prd HD K9 and Gsc HD. (D) Cooperative heterodimerization of the Gsc and Mix-1 HDs on P3C. The positions of the singly (M) or doubly (D) bound probes by either the Gsc HD or the Mix-1 HD are marked. The two cooperative homodimers (Gsc D and Mix D), and the cooperative heterodimer (Gsc/Mix D) are formed on P3C at these protein concentrations but are very faint on P5C.

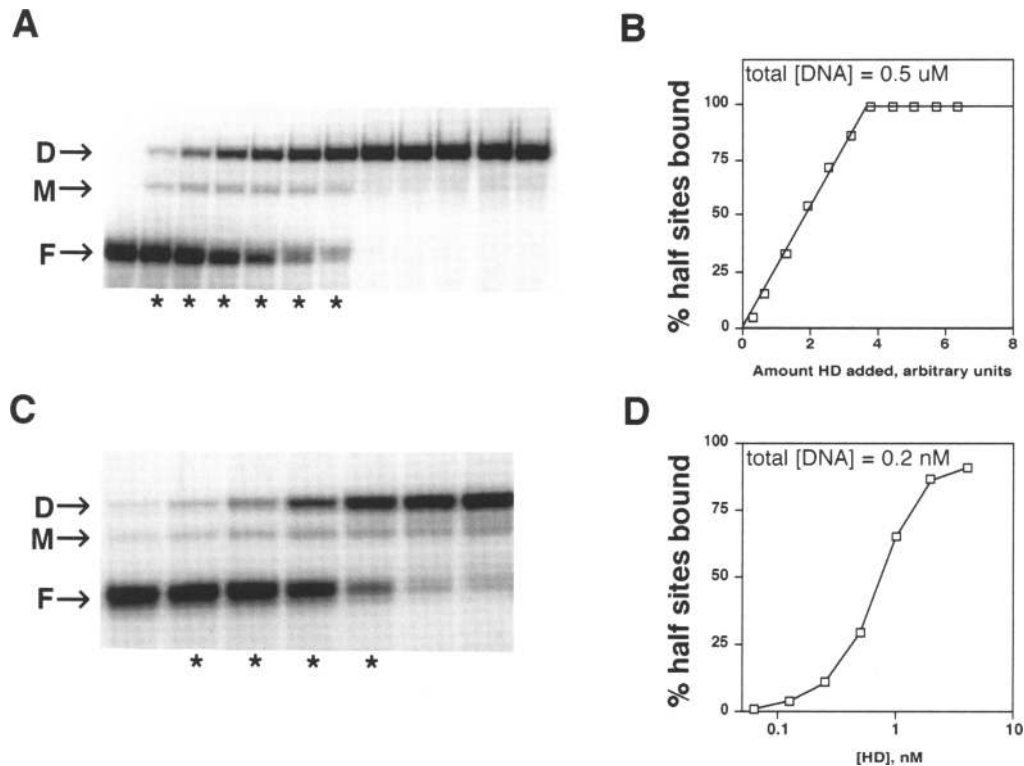
1987). Surprisingly, the Prd HD binds optimally to a palindromic site composed of two inverted TAAT motifs, which are separated by 2 bp (P2). This is in sharp contrast to the monomeric mode with which other HDs are believed to recognize DNA (Pabo and Sauer 1992). The 60 amino acids comprising the Prd HD are sufficient to engage in cooperative binding to the two half-sites in this arrangement: the first binding event stimulates the second one by ~50-fold. This cooperativity is lost when the inverted TAATs are disjointed, that is, separated by 5 bp (P5). In *Drosophila* cell culture, the full-length Prd protein activates transcription considerably more effectively (40-fold) on a single P2 site than on a P5 site or a half-site, suggesting that it can dimerize cooperatively on the selected P2 site in vivo.

The cooperative binding potential of Prd is conserved, as other HDs in the Prd class, such as Gsb or the more divergent known members, Lune, Mix-1, or Gsc, also bind cooperatively to the selected palindromic se-

quences (Fig. 8). Furthermore, cooperative heterodimers between HDs in this class can also form on the palindromic binding sites. Because Prd and Gsb are expressed in the same *Drosophila* cells at the same time during embryogenesis (Frigerio et al. 1986; Gutjahr et al. 1993), the Gsb-Prd cooperative heterodimer (Fig. 2B) may be of regulatory importance. Consistent with this model, Morrissey et al. (1991) have reported a transdominant effect of an inactive version of Prd on *gsb* function: By dimerizing with Gsb on DNA, an inactive Prd protein may obliterate *gsb* function without altering its expression pattern. In the mouse, overlapping expression patterns of Prd class HD proteins (Pax genes; Gruss and Walther 1992) also suggest the formation of heterodimeric species. Furthermore, the overlapping expression patterns of the *Xenopus* homeo proteins Mix-1 and Gsc may result in the formation of the cooperative DNA-binding heterodimer shown in Figure 6D. The possibility of heterotypic combinations of Prd class HDs



Wilson et al.



**Figure 7.** Quantification of binding cooperativity and dissociation constants for Prd HD Q9 binding to the P3 site. (A) Titration of P3 probe with increasing amounts of purified Prd HD Q9 peptide to determine HD concentration. The total probe concentration is 0.5  $\mu\text{M}$ . (B) Graph of probe binding from mobility shift shown in A, which demonstrates that the DNA concentration used in this study is very far above the dissociation constants for interaction between the Prd HD Q9 and the P3 site. (C) Mobility shift of P3 at low probe concentration (0.2 nM) by Prd HD Q9 in the concentration range of 62.5  $\mu\text{M}$  to 4 nM. (D) Graph of probe binding from mobility shift shown in C, showing DNA binding to be limited by affinity per se rather than merely by protein amount (to contrast with the titration shown in A). The asterisks are referred to in the text.

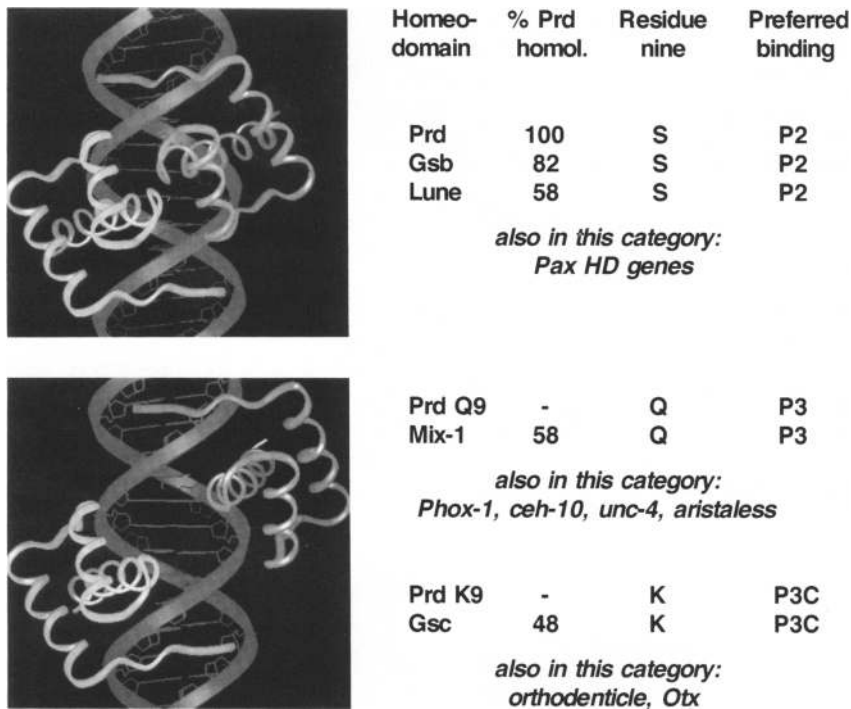
may multiply the number of homeo protein species capable of executing distinct functions.

Although we could not detect cooperative dimerization on DNA of HDs in other classes (Engrailed, Even-skipped, Ftz, Zerknüllt, S59, Distalless; for references, see Bürglin 1993), it remains possible that members of these classes can contribute to heterodimers with members of other classes in combinations not yet identified.

#### Specificity of HD dimerization

The ninth amino acid position of the recognition helix has been shown previously to interact with the base pairs immediately 3' to the TAAT core sequence (Hanes and Brent 1989, 1991; Percival-Smith et al. 1990; Treisman et al. 1992). These base pairs correspond to the center of symmetry in the palindromic sequences recognized by Prd class HDs. In relation to this, we show that the ninth amino acid position can change the nature of the cooperative complex formed by Prd class HDs. Whereas Prd class members possessing a serine at position 9 (S9) bind cooperatively to a palindromic DNA site in which 2 bp separate the two TAATs (P2), those car-

rying a glutamine at this position (Q9) prefer 3 bp between the TAATs (P3). The Q9-type HDs possess a stronger cooperativity than the S9 members, where the first HD increases the affinity of the second by >200-fold on P3, as compared with ~50-fold in the case of the Prd HD on P2. Those HDs in which the position is occupied by a lysine also prefer 3 bp between the inverted TAAT motifs for cooperative interaction, but they have a more stringent requirement for the identity of the base pairs between the TAAT motifs and display a weaker cooperativity (25-fold). In the context of the Prd class of HDs, then, the ninth amino acid of the recognition helix simultaneously performs three functions: to set the specificity of spacing between TAAT half-sites; to modulate the magnitude of the cooperative interaction; and to determine the nature of the base pairs in the center of the recognition sequence. Hence, the Prd HD class includes members with at least three distinct DNA-binding specificities attributable to the identity of one residue. Other studies have attempted to distinguish the DNA-binding preferences between HDs within the Antennapedia class (Florence et al. 1991; Dessain et al. 1992; Ekker et al. 1992). The differences between these members in DNA-binding specificity are more subtle, possibly because



**Figure 8.** The three recognition specificities of Prd class HDs. Models of the conformations for optimal binding complexes of Prd class HDs containing a serine (*top*) vs. a glutamine or a lysine (*bottom*) at recognition helix position 9. Models are based on coordinates kindly supplied by Carl Pabo for the Engrailed HD binding to a TAATTA sequence (Kissinger et al. 1990). For the HDs tested in this study, the preferred binding sites, homology to Prd HD, and residues at recognition helix position 9 are shown. For the three categories of Prd class HDs (defined by the amino acid at position nine, S, Q, or K), other HDs are indicated and are predicted to have similar DNA-binding specificities as the members tested in the same category. All HDs referred to have been classified recently (Bürglin 1993).

they all possess the same residue (glutamine) at recognition helix position 9.

No natural target sequences for Prd class homeo proteins have yet been defined. A 1-kb enhancer region from the *engrailed* gene, which directs *prd*-dependent expression (Kassis 1990), however, contains the site TAATT-TAATTG, which matches 9 of 10 bp with P3 and is conserved between *Drosophila melanogaster* and *D. virilis* (Kassis et al. 1989).

#### Structures of Prd class HD cooperative dimers

Most animal homeo proteins are believed to bind to single TAAT-containing sites, or to more complex sequences as a result of interaction with other protein domains. It seems likely (see next paragraph) that in the palindromic sequences, the two Prd HD molecules dock symmetrically onto each half-site with a conformation similar to that described for HDs binding individual TAAT motifs (Kissinger et al. 1990; Otting et al. 1990). A model of the P2 complex, based on the orientation of the Engrailed HD binding to a single TAAT site (coordinates provided by C. Pabo), is shown in Figure 8. The two recognition helices approach each other at their amino termini and could potentially interact through the turn of the helix-turn-helix to mediate the cooperative interaction.

The conversion from S to Q or K at position 9 of the recognition helix results in a dramatic structural change of the cooperative complex on P2 versus P3 (Fig. 8): The 1 bp increased spacing results in a vertical and rotational shift of the two HDs away from each other (by  $\sim 3.4$  Å and  $34^\circ$ , respectively), as shown in the comparison be-

tween the top (P2) and bottom (P3) structures in Figure 8. Modeling of the P3 dimer suggests that this structural change (from P2 to P3) would prevent any contact between the two HD monomers and, therefore, disable the cooperative interaction between them. In contrast, increased cooperativity is observed with the Q9 (P3) spacing. Three possible models may explain this paradox. (1) The DNA may bend in the P3 complex, as is suggested by the aberrant migration of the cooperative dimer on the P3 probe (see Fig. 4B); however, a very severe deformation would be required to maintain contacts similar to those present in the P2 complex. (2) Alternatively, the structural features of the individual Prd HDs binding to the P3 half-sites may be different from those defined by X-ray crystallography and nuclear magnetic resonance (NMR) for other HDs. More divergent HDs (Antennapedia and  $\alpha 2$ , which share only 24% homology), however, have extremely related structures and DNA docking conformations (Otting et al. 1990; Wolberger et al. 1991). Furthermore, because both Antennapedia class (Nelson and Laughon 1990; Ekker et al. 1991, 1992; Florence et al. 1991) and Prd class HDs prefer binding to sequences containing a TAAT motif to all other random sequences, similar amino acid-base pair contacts have probably been maintained throughout evolution, thus eliminating the likelihood of major structural differences. (3) Finally, the cooperative effects may be mediated through an allosteric change in the DNA; that is, the first binding event to one half-site may change the conformation of the second half-site, improving its affinity for a second HD. Such a mechanism would not require contact between the two HDs. DNA has been reported capable of mediating allostery (Hogan et al. 1979; Ansari et al.

Wilson et al.

1992) but not with such spatial specificity as to allow cooperativity differences on, for example, P3 versus P5.

Examination of the amino acid sequences specifically conserved among the Prd class HDs shown to engage in cooperative dimerization suggest possible regions responsible for the interaction. The loop between the first and second helices contains the conserved sequence YPXI/V at positions 25–28 (*X* is a nonconserved position). This sequence in one HD may contact the relatively nearby basic amino acids at the amino-terminal arm of the second HD, which are highly conserved among all HDs and interact with the minor groove of the DNA. Also specifically conserved among the Prd class HDs are recognition helix positions one (E) and three (R or K), which may interact electrostatically with each other in the cooperative complex. The residues of helix 2 at positions 41–44, REXL, are also specifically conserved among the HDs shown to dimerize cooperatively, although the mechanism by which they might contribute to this effect is beyond prediction.

#### *HD dimerization as a functional specificity determinant*

The high degree of conservation of the HD and previous reports of its modest DNA-binding specificity have suggested that functional specificity differences between different homeo proteins could not derive merely from inherent differences in DNA-binding preference (Hayashi and Scott 1990). Some homeo proteins therefore rely on protein sequence outside of HD for specificity; examples are those homeo proteins bearing a second DNA-binding domain, such as the Prd domain or the POU-specific domain (Sturm and Herr 1988; Ingraham et al. 1990; Chalepakis et al. 1991; Treisman et al. 1991; Verrijzer et al. 1992). Homeo proteins have also achieved specificity by requiring a cofactor to bind DNA efficiently (Mendel et al. 1991). A well-documented example is the yeast  $\alpha 2$  protein, able to interact with the MCM-1 factor (Keleher et al. 1988; Smith and Johnson 1992). Some homeo proteins, such as HNF-1 and  $\alpha 2$ , contain independent dimerization domains that allow the formation of homo- and heterodimers (Dranginis 1990; Rey-Campos et al. 1991). The POU-specific domain can mediate DNA-dependent dimerization (Ingraham et al. 1990), which may explain the observation that the intact Oct-1 POU domain (POU-specific domain plus POU HD) can dimerize cooperatively on adjacent octamer sites (LeBowitz et al. 1989). It should be noted that in this case the requirements for orientation and spacing of the two cooperating binding sites for Oct-1 (LeBowitz et al. 1989) are completely different from those for the HD-mediated cooperativity reported here for the Prd class HDs, as judged from a recent alignment of the TAAT motif with the octamer sequence (Verrijzer et al. 1992).

In contrast to this reliance on protein elements outside of the HD, some HDs themselves contain sufficient information to provide high level specificity. For instance, a Prd class HD has been reported recently to interact with members of the SRF/MCM-1 family of transcrip-

tion factors (Grueneberg et al. 1992). Also, the Oct-1 HD can interact with the VP16 protein (Lai et al. 1992; Pomerantz et al. 1992), and some POU HDs have been shown recently to interact with each other in solution (Treacy et al. 1991, 1992; Voss et al. 1991). Here, we report a new strategy by which a property intrinsic to the HD can provide increased specificity: The Prd class HDs contain sufficient information to specify the assembly of cooperative homo- and heterodimers on palindromic DNA sequences. These palindromic binding sites are longer and therefore occur more rarely in the genome than the sequences recognized by HDs in other classes, thus increasing the binding specificity. Our observation may be related to the finding that the isolated POU HD, which shares only 29% identity with Prd, can dimerize on a sequence that resembles the P3 site (TAATGAPuATT vs. TAATPyNPuATTA; Kristie and Sharp 1990). Although the cooperative binding by the Oct-1 HD is apparently abolished by the POU-specific domain, these results suggest that a wide variety of HDs have the potential to dimerize on DNA.

The ability of a small globular domain to recognize distinct configurations of binding sites on DNA has also recently been reported for other DNA-binding domains. Different members of the bZIP family (Hai et al. 1988; Sellers et al. 1990; Cuenoud and Schepartz 1993) and the  $C_6$  zinc cluster family (Reece and Ptashne 1993) recognize palindromes with different spacings between half-sites. The same is true for members of the steroid receptor superfamily (Umesono and Evans 1989; Luisi et al. 1991), some members of which instead prefer binding to direct repeats with different spacings (Umesono et al. 1991) or even to monomeric sites (Wilson et al. 1992). The mechanisms by which different members of these conserved families have evolved distinct configurations for DNA recognition are of considerable interest.

## Materials and methods

### *Cloning and protein preparation*

HD peptides used for mobility shifts were produced by cloning homeo box sequences into the *NdeI*-*EcoRI* sites of pAR3038, overexpressing them in *E. coli* strain BL21, and preparing them as crude extracts as described (Treisman et al. 1989). The restriction sites were introduced into the homeo boxes by PCR. Wild-type and mutant Prd HD peptides consisted of an introduced initiating methionine, the 60-amino-acid HD, and 4 amino acids downstream derived from Prd. Mutagenesis at position 9 was performed with a 3' PCR primer degenerate at this codon. The Gsb HD has 19 amino acids upstream and 9 amino acids downstream of the HD. The Mix-1 HD has 17 amino acids upstream and 4 amino acids downstream of the HD. The Prd HD Q9 peptide was purified to >85% purity by FPLC on heparin followed by Mono-S (Pharmacia) and will be described elsewhere (B. Guenther, unpubl.). The full-length Mix-1 protein was cloned into the same vector using the *NdeI* and *BamHI* sites and expressed in the coupled in vitro transcription/translation rabbit reticulocyte system (Promega) with the T7 polymerase. Equal amounts (1  $\mu$ g) of either pAR Mix-1 or the control vector without insert were added to a 50  $\mu$ l Promega coupled TnT transcription/translation reaction. All protein preparations were stored at  $-80^\circ\text{C}$ .



GST fusion proteins were created in the pGEX-2T vector (Pharmacia) after introducing *Bam*HI and *Eco*RI sites into homeo box inserts by PCR. For the wild-type and mutant Prd HDs, 17 amino acids upstream and 4 amino acids downstream of the HD (derived from Prd) were included in the expressed product. The proteins were expressed and purified as described (Ausubel et al. 1991). Protein was not eluted from the beads. Protein/bead slurries are stable for several months at 4°C.

The plasmid producing Prd was cloned by inserting *Nde*I-cut/blunt-ended cDNA sequence into the blunt-ended unique *Bam*HI site in the pPAC vector (Jaynes and O'Farrell 1988; Krasnow et al. 1989). The putative HD-binding sequences were cloned as follows: A linker containing a *Bam*HI site was cloned into the unique *Pst*I site of the pD-33CAT vector (Ronchi et al. 1993). The annealed oligonucleotides described above for the mobility shift (P1/2, P2, and P5<sub>2</sub>) were then cloned into the *Bam*HI site of the modified pD-33CAT vector.

#### Random selection protocol

The random sequence library was prepared by synthesizing an oligonucleotide with the sequence CATGAATTCTCCTATACTGAGTTTCATGATN18TGATATCGAACTGTATCGATGATTCCAC (or related oligonucleotides) and PCR primers annealing to the first (top strand) and last (bottom strand) 20 bases. The 3' primer was gel-purified and end-labeled and a primer extension reaction was carried out with purified random template. The selection procedure was initiated by mixing 100 ng of the double-stranded template with 10 ng of fusion protein attached to 5 µl of GSH-agarose beads in 0.5 ml of 20 mM Tris (pH 8.0), 50 mM KCl, 0.5 mM EDTA, 1 mM DTT, 10% glycerol, 20 µg/ml of BSA, and 2 µg/ml of poly[d(I-C)]. The mixture was rotated at 4°C for 1 hr and then was centrifuged at 12,000g for 1 min. The supernatant was removed, and the pellet was washed twice with 0.8 ml of 20 mM Tris-HCl (pH 8.0), 50 mM KCl, 0.5 mM EDTA, 1 mM DTT, and 10% glycerol. This pellet was resuspended in 30 µl of water, boiled for 3 min, centrifuged rapidly, and 10 µl of this supernatant was used as template for a PCR reaction containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 200 µM each dNTP, 10 µg/ml of each primer, and 50 U/ml of *Taq* polymerase in a total volume of 50 µl. Twenty cycles of 94°C (30 sec), 44°C (30 sec), and 72°C (30 sec) were followed by the addition of excess amounts (1.2 gram) of each primer and 2.5 additional units of *Taq* polymerase. A final cycle of 94°C (2 min), 44°C (1 min), and 72°C (10 min) ensured that most of the library would be double stranded without mismatches. This would conclude the first generation of selection/amplification. The second generation begins when 10 µl of the above PCR reaction (which contains 100 ng of double-stranded random DNA template) is mixed with fusion protein exactly as in the first generation. The remaining steps are identical to that for the first generation. After the last PCR amplification step, a fraction of the PCR product is cut with *Eco*RI and cloned into PKSII- (Stratagene), and the individual selected oligonucleotides are sequenced. Selections consisted of five generations, except where noted.

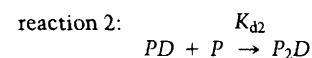
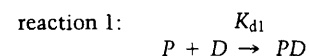
#### Mobility shift assay

Oligonucleotides used for the mobility shift assay were purified on sequencing gels, eluted into 0.5 M ammonium acetate, 10 mM Tris-HCl, 1 mM EDTA, 0.1% SDS (pH 5.2), filtered, and desalted on a NAP-25 Sephadex column. The two strands of each probe possess 5' GATC overhangs to permit labeling with DNA polymerase, cold dGTP, dTTP, and dCTP, and [ $\alpha$ -<sup>32</sup>P]dATP. The sequences for the top strand of the probes

are as follows: P2, GATCCTGAGTCTAATTGATTACTGTAC-AG; P5<sub>2</sub>, GATCTGAGTCTAATTGATGATTACTGTAC; P3, GATCCTGAGTCTAATTGAATTACTGTACA; P5<sub>3</sub>, GATCTGAGTCTAATTGAGAATTACTGTAC; P3C, GATCCTGAGTCTAATCCGATTACTGTACA; P5C, GATCTGAGTCTAATCCGATTACTGTAC; and P1/2, GATCCTGAGTCTAATTGAGCGTCTGTACA. For the P2 probe, the sequence for the bottom strand is GATCCTGTACAGTAATCAATTAGACTCAG, and analogous bottom strands were used for the remaining probes. For the mobility shift shown in Figure 7A, the probe was top strand, AGACTTAATTGAATTATCAGC, and bottom strand, TGCTGATAATTCAATTAAGTC. This probe was labeled by kinase with [ $\gamma$ -<sup>32</sup>P]dATP and then annealed. The crude protein extracts were diluted in 20 mM Tris-HCl (pH 7.5), 75 mM KCl, 1 mM DTT, 0.5 mg/ml of BSA, 1 mM PMSF, 200 µM benzamidine, and 10% glycerol to appropriate levels where binding to the indicated probes could be detected, but no binding was observed in a mock extract or to nonspecific DNA. Diluted extracts were mixed with 0.2 ng of labeled DNA (except for Fig. 7, as indicated) in a total volume of 20 µl containing 15 mM Tris-HCl (pH 7.5), 60 mM KCl, 0.5 mM DTT, 0.25 mg/ml of BSA, 0.5 mM PMSF, 100 µM benzamidine, 4 mM spermine, 4 mM spermidine, 50 µg/ml of poly[d(I-C)], 0.05% NP-40, and 7.5% glycerol. For the mobility shifts with purified protein, protein dilutions and binding reactions omitted PMSF, benzamidine, spermine, spermidine, and poly[d(I-C)]. After 25 min at room temperature, contents were loaded onto an 8% polyacrylamide/bis-acrylamide (29 : 1) gel buffered with 0.25× TBE, and run at 12 V/cm in 0.25× TBE running buffer for 2.5 hr. Gels were then fixed in 10% acetic acid, 10% ethanol, dried, and exposed to film or PhosphorImager cassette. For the experiment with the full-length proteins transcribed/translated in vitro, the reactions were run on a 4% gel for 1.5 hr. Quantification of mobility shift bands was performed on the PhosphorImager using the integrate volume function. Estimation of Prd HD and Prd HD Q9 concentration in the crude lysates was made by comparing purified Prd HD Q9 peptide, at a concentration determined in Figure 7B, with crude lysates on a SDS-polyacrylamide gel. This was used to estimate the concentration of HD in the gels shown in Figures 2A and 4B and is accurate ±50%.

#### Calculations of DNA-binding affinity and cooperativity

The binding of a monomeric protein to a DNA probe containing two possible binding sites is represented by



where *P* is the protein, *D* is the DNA, *PD* is the singly occupied DNA, and *P<sub>2</sub>D* is the dimerically bound DNA. *K<sub>d1</sub>* and *K<sub>d2</sub>* are the dissociation constants for the reverse of the first and second reaction, respectively, and are equal to

$$K_{d1} = [PD]/2[P][D] \quad (1)$$

$$K_{d2} = 2[P_2D]/[PD][P] \quad (2)$$

(the 2's are statistical factors). Here, [*P*] is the free, rather than the total, protein concentration. The cooperativity factor,  $\tau$ , is defined as *K<sub>d2</sub>*/*K<sub>d1</sub>*, that is, the extent to which one bound protein aids a second protein in binding. Therefore,

$$\tau = 4[P_2D][D]/[PD]^2 \quad (3)$$

Wilson et al.

$\tau$  is unitless and should be independent of the concentration of components. Equations 1 and 2 can be combined to give

$$K_{d2} = [P_2D]/K_{d1}[D][P]^2 \quad (4)$$

When two probes are compared that have the same pair of half-sites (e.g., P3 vs. P5<sub>3</sub>) but differ in the strength of cooperativity, the only difference between them should be in the value of  $K_{d2}$ . The ratio of  $K_{d2}$  for the noncooperative probe over the cooperative probe is therefore equal to  $\tau$ . Equation 4 shows that the ratio of protein concentrations required to shift a noncooperative versus a cooperative probe to the point where  $[P_2D] = [D]$  is equal to the square root of  $\tau$ . The equations used in determining  $K_{d1}$ ,  $K_{d2}$ , and  $\tau$  are mentioned in Results. When shown, errors are expressed as the standard error of the mean.

#### *Drosophila cell culture and CAT assays*

*Drosophila* Schneider 2 cells were cultured in M3 media supplemented with 10% fetal bovine serum at 25°C. Transfection was done at 50–70% confluency by the calcium phosphate procedure. A total of 15  $\mu$ g of DNA containing 1  $\mu$ g of *hsp82*- $\beta$ -gal and 2  $\mu$ g of reporter constructs were used for each transfection. Cells were then grown for 36–48 hr, after which they were washed three times with PBS, once with TEN (40 mM Tris-HCl at pH 7.5, 1 mM EDTA, 150 mM NaCl), lysed in 150  $\mu$ l of 0.25 M Tris-HCl (pH 7.5) by freeze-thawing three times, and supernatant was collected for assay. Chloramphenicol acetyltransferase (CAT) assays were carried out by mixing 150  $\mu$ l of 0.25 M Tris-HCl (pH 7.5), 20  $\mu$ l of the extract, 10  $\mu$ l of 4 mM AcCoA (Boehringer), and 3  $\mu$ l of [<sup>14</sup>C]chloramphenicol (DuPont), and incubating for 1 hr at 37°C. The reaction was stopped by extraction with ethyl acetate. The extract was lyophilized and spotted on TLC plate (Eastman Chromogram) and run in 19 : 1 chloroform/methanol. Radiolabeled chloramphenicol and acetylated chloramphenicol were quantified with a PhosphorImager using the integrate volume function. CAT activities were normalized according to the  $\beta$ -gal activity.

#### Acknowledgments

D.W. is supported by a National Research Service award (GM-07982-12). N.D. is a recipient of a Human Frontiers Fellowship. Brian Guenther, Michael Stella, X.-P. Kong, and Dorothy Kominos created the graphic models in Figure 8 using coordinates kindly provided by C. Pabo. We thank Susie Jun and Bill Kalionis for supplying and communicating unpublished results concerning the Lune gene. B. Guenther purified the Prd HD Q9 peptide in the laboratory of John Kuriyan, The Rockefeller University. Jan Geliebter and Chris Marshall provided excellent technical advice. The cell culture Prd producer construct was made by Elettra Ronchi. Jessica Treisman was instrumental in the genesis of this work. Thanks go to Yan Gu for excellent technical assistance, and to Susie Jun, Pierre Gönczy, Scott Dougan, Steve DiNardo, Carl Pabo, and Titia de Lange for critical reading of the manuscript.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

#### References

- Affolter, M., A. Percival-Smith, M. Muller, W. Leupin, and W.J. Gehring. 1991. DNA binding properties of the purified Antennapedia homeodomain. *Proc. Natl. Acad. Sci.* **87**: 4093–4097.
- Ansari, A.Z., M. Chael, and T.V. O'Halloran. 1992. Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature* **355**: 87–89.
- Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, and J.A. Smith, eds. 1990. *Current protocols in molecular biology*. John Wiley and Sons, New York.
- Baumgartner, S., D. Bopp, M. Burri, and M. Noll. 1987. Structure of two genes at the gooseberry locus related to the paired gene and their spatial expression during *Drosophila* embryogenesis. *Genes & Dev.* **1**: 1247–1267.
- Blumberg, B., C. Wright, E. DeRobertis, and K. Cho. 1991. Organizer-specific homeobox genes in *Xenopus laevis* embryos. *Science* **253**: 194–196.
- Bopp, D., M. Burri, S. Baumgartner, G. Frigerio, and M. Noll. 1986. Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes in *Drosophila*. *Cell* **47**: 1033–1049.
- Bürglin, T.R. 1993. A comprehensive classification of homeobox genes. In *A guidebook for homeobox genes* (ed. D. Duboule). Oxford University Press, England. (In press.)
- Chalepakis, G., R. Fritsch, H. Fickenscher, U. Deutsch, M. Goulding, and P. Gruss. 1991. The molecular basis of the undulated/Pax-1 mutation. *Cell* **66**: 873–884.
- Cuenoud, B. and A. Schepartz. 1993. Altered specificity of DNA-binding proteins with transition metal dimerization domains. *Science* **259**: 510–513.
- Desplan, C., J. Theis and P.H. O'Farrell. 1988. The sequence specificity of homeodomain-DNA interaction. *Cell* **54**: 1081–1090.
- Dessain, S., C.T. Gross, M.A. Kuziora, and W. McGinnis. 1992. Antp-type homeodomains have distinct DNA binding specificities that correlate with their regulatory functions in embryos. *EMBO J.* **11**: 991–1002.
- Dranginis, A.M. 1990. Binding of yeast  $\alpha 1$  and  $\alpha 2$  as a heterodimer to the operator DNA of a haploid-specific gene. *Nature* **347**: 682–685.
- Driever, W. and C. Nüsslein-Volhard. 1989. The bicoid protein is a positive regulator of *hunchback* transcription in the early *Drosophila* embryo. *Nature* **337**: 138–143.
- Ekker, S.C., K.E. Young, D.P. von Kessler, and P.A. Beachy. 1991. Optimal DNA sequence recognition by the ultrabithorax homeodomain of *Drosophila*. *EMBO J.* **10**: 1179–1186.
- Ekker, S.C., D.P. von Kessler, and P.A. Beachy. 1992. Differential DNA sequence recognition is a determinant of specificity in homeotic gene action. *EMBO J.* **11**: 4059–4072.
- Epstein, D.J., M. Vekemans, and P. Gros. 1991. *splotch* (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell* **67**: 767–774.
- Finkelstein, R., D. Smouse, T.M. Capaci, A.C. Spradling, and N. Perrimon. 1990. The *orthodenticle* gene encodes a novel homeo domain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures. *Genes & Dev.* **4**: 1516–1527.
- Florence, B., R. Handrow, and A. Laughon. 1991. DNA-binding specificity of the fushi tarazu homeodomain. *Mol. Cell. Biol.* **11**: 3613–3623.
- Fortini, M.E., Z. Lai, and G.M. Rubin. 1991. The *Drosophila* *zfh-1* and *zfh-2* genes encode novel proteins containing both zinc-finger and homeodomain motifs. *Mech. Dev.* **34**: 113–122.
- Freemont, P.S., A.N. Lane, and M.R. Sanderson. 1991. Structural aspects of protein-DNA recognition. *Biochem. J.* **278**: 1–23.

- Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation of gene paired and the *Drosophila* PRD gene set as part of a gene network. *Cell* **47**: 735–746.
- Furukubo-Tokunaga, K., M. Muller, M. Affolter, L. Pick, U. Kloter, and W.J. Gehring. 1992. In vivo analysis of the helix–turn–helix motif of the *fushi tarazu* homeo domain of *Drosophila melanogaster*. *Genes & Dev.* **6**: 1082–1096.
- Gibson, G., A. Schier, P. LeMotte, and W.J. Gehring. 1990. The specificities of sex combs reduced and antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell* **62**: 1087–1103.
- Grueneberg, D.A., S. Natesan, C. Alexandre, and M.Z. Gilman. 1992. Human and *Drosophila* homeodomain proteins that enhance the DNA-binding activity of serum response factor. *Science* **257**: 1089–1095.
- Gruss, P. and C. Walther. 1992. Pax in development. *Cell* **69**: 719–772.
- Gutjahr, T., I. Frei, and M. Noll. 1993. Complex regulation of early paired expression: Initial activation by gap genes and pattern modulation by pair-rule genes. *Development* **117**: 609–623.
- Hai, T., F. Liu, E.A. Allegretto, M. Karin, and M.R. Green. 1988. A family of immunologically related transcription factors that includes multiple forms of ATF and AP-1. *Genes & Dev.* **2**: 1216–1226.
- Hanes, S.D. and R. Brent. 1989. DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. *Cell* **57**: 1275–1283.
- . 1991. A genetic model for interaction of the homeodomain recognition helix with DNA. *Science* **251**: 426–430.
- Hayashi, S. and M. Scott. 1990. What determines the specificity of action of *Drosophila* homeodomain proteins? *Cell* **63**: 883–894.
- Hill, R.E., J. Favor, B.L.M. Hogan, C.C.T. Ton, G.F. Saunders, L.M. Hanson, J. Prosser, T. Jordan, N.D. Hastie, and B. van Heyningen. 1991. Mouse Small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354**: 522–525.
- Hoey, T. and M. Levine. 1988. Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. *Nature* **332**: 858–861.
- Hogan, M., N. Dattagupta, and D.M. Crothers. 1979. Transmission of allosteric effects in DNA. *Nature* **278**: 521–524.
- Hope, I.A. and K. Struhl. 1987. GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.* **6**: 2781–2784.
- Ingraham, H.A., S.E. Flynn, J.W. Voss, V.R. Albert, M.S. Kapiloff, L. Wilson, and M.G. Rosenfeld. 1990. The POU-specific domain of Pit-1 is essential for sequence-specific high affinity DNA binding and DNA-dependent Pit-1-Pit-1 dimerizations. *Cell* **61**: 1021–1033.
- Jaynes, J.B. and P.H. O'Farrell. 1988. Activation and repression of transcription by homeodomain-containing proteins that bind a common site. *Nature* **336**: 744–749.
- Kassis, J.A. 1990. Spatial and temporal control elements of the *Drosophila engrailed* gene. *Genes & Dev.* **4**: 433–443.
- Kassis, J.A., C. Desplan, D.K. Wright, and P.H. O'Farrell. 1989. Evolutionary conservation of homeodomain-binding sites and other sequences upstream and within the major transcription unit of the *Drosophila* segmentation gene *engrailed*. *Mol. Cell. Biol.* **9**: 4304–4311.
- Keleher, C.A., C. Goutte, and A.D. Johnson. 1988. The yeast cell-type-specific repressor  $\alpha 2$  acts cooperatively with a non-cell-type-specific protein. *Cell* **53**: 927–936.
- Kissinger, C.R., B. Liu, E. Martin-Blanco, T.B. Kornberg, and C.O. Pabo. 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 angstrom resolution: A framework for understanding homeodomain-DNA interactions. *Cell* **63**: 579–590.
- Krasnow, M.A., E.E. Saffman, K. Kornfeld, and D.S. Hogness. 1989. Transcriptional activation and repression by ultrabithorax proteins in cultured *Drosophila* cells. *Cell* **57**: 1031–1043.
- Kristie, T.M. and P.A. Sharp. 1990. Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV  $\alpha$ -trans-activator protein. *Genes & Dev.* **4**: 2383–2396.
- Kuziora, M.A. and W. McGinnis. 1989. A homeodomain substitution changes the regulatory specificity of the deformed protein in *Drosophila* embryos. *Cell* **59**: 563–571.
- Lai, J.-S., M.A. Cleary, and W. Herr. 1992. A single amino acid exchange transfers VP16-induced positive control from the Oct-1 to the Oct-2 homeo domain. *Genes & Dev.* **6**: 2058–2065.
- Laughon, A. and M.P. Scott. 1984. Sequence of a *Drosophila* segmentation gene: Protein structure homology with DNA binding proteins. *Nature* **310**: 25–31.
- LeBowitz, J.H., R.G. Clerc, M. Brenowitz, and P.A. Sharp. 1989. The Oct-2 protein binds cooperatively to adjacent octamer sites. *Genes & Dev.* **3**: 1625–1638.
- Luisi, B.F., W.X. Wu, Z. Otwinowski, L.P. Freedman, K.R. Yamamoto, and P.B. Sigler. 1991. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* **352**: 497–505.
- Mann, R.S. and D.S. Hogness. 1990. Functional dissection of ultrabithorax proteins in *D. melanogaster*. *Cell* **60**: 597–610.
- McGinnis, W. and R. Krumlauf. 1992. Homeobox genes and axial patterning. *Cell* **68**: 283–302.
- Mendel, D.B., P.A. Khavari, P.B. Conley, M.K. Graves, L.P. Hansen, A. Admon, and G.R. Crabtree. 1991. Characterization of a cofactor that regulates dimerization of a mammalian homeodomain protein. *Science* **254**: 1762–1767.
- Miller, D.M., M.M. Shen, C.E. Shamu, T.R. Burglin, G. Ruvkun, M.L. Dubois, M. Ghee, and L. Wilson. 1992. *C. elegans unc-4* gene encodes a homeodomain protein that determines the pattern of synaptic input to specific motor neurons. *Nature* **335**: 841–845.
- Morrissey, D., D. Askew, L. Raj, and M. Weir. 1991. Functional dissection of the *paired* segmentation gene in *Drosophila* embryos. *Genes & Dev.* **5**: 1684–1696.
- Nelson, H.B. and A. Laughon. 1990. The DNA binding specificity of the *Drosophila fushi tarazu* protein: A possible role for DNA bending in homeodomain recognition. *New Biol.* **2**: 171–178.
- Otting, G., Y.Q. Qian, M. Billeter, M. Muller, M. Affolter, W.J. Gehring, and K. Wüthrich. 1990. Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* **9**: 3085–3092.
- Pabo, C.O. and R.T. Sauer. 1992. Transcription factors: Structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**: 1053–1095.
- Percival-Smith, A., M. Muller, M. Affolter, and W. Gehring. 1990. The interaction with DNA of wild-type and mutant *fushi tarazu* homeodomains. *EMBO J.* **9**: 3967–3974.
- Pomerantz, J.L., T.M. Kristie, and P.A. Sharp. 1992. Recognition of the surface of a homeo domain protein. *Genes & Dev.* **6**: 2047–2057.
- Qian, Y.Q., M. Billeter, G. Otting, M. Muller, W.J. Gehring, and K. Wüthrich. 1989. The structure of the antennapedia homeo-



- odomain determined by NMR spectroscopy in solution: Comparison with prokaryotic repressors. *Cell* **59**: 573–580.
- Reece, R.J. and M. Ptashne. 1993. Determinants of binding-site specificity among yeast C6 zinc cluster proteins. *Science* **61**: 909–911.
- Rey-Campos, J., T. Chouard, M. Yaniv, and S. Cereghini. 1991. vHNF1 is a homeoprotein that activates transcription and forms heterodimers with HNF1. *EMBO J.* **10**: 1445–1457.
- Rosa, F.M. 1989. Mix.1 a homeobox mRNA inducible by mesoderm inducers is expressed mostly in the presumptive endodermal cells of *Xenopus* embryos. *Cell* **57**: 965–974.
- Ronchi, E., J. Treisman, N. Dostatni, G. Struhl, and C. Desplan. 1993. Down-regulation of the *Drosophila* morphogen Bicoid by the Torso receptor-mediated signal transduction cascade. *Cell* **74**: 347–355.
- Ruberti, I., G. Sessa, S. Lucchetti, and G. Morelli. 1991. A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. *EMBO J.* **10**: 1787–1791.
- Schier, A.F. and W.J. Gehring. 1992. Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene. *Nature* **356**: 804–807.
- . 1993. Functional specificity of the homeodomain protein *fushi tarazu*: The role of DNA-binding specificity in vivo. *Proc. Natl. Acad. Sci.* **90**: 1450–1454.
- Scott, M.P., J.W. Tamkun, and G.W. Hartzell. 1989. The structure and function of the homeodomain. *Biochem. Biophys. Acta* **989**: 25–48.
- Sellers, J.W., A.C. Vincent, and K. Struhl. 1990. Mutations that define the optimal half-site for binding yeast GCN4 activator protein and identify an ATF/CREB-like repressor that recognizes similar DNA sites. *Mol. Cell Biol.* **10**: 5077–5086.
- Smith, D.L. and A.D. Johnson. 1992. A molecular mechanism for combinatorial control in yeast: MCM1 protein sets the spacing and orientation of homeodomains of an  $\alpha 2$  dimer. *Cell* **68**: 133–142.
- Sturm, R.A. and W. Herr. 1988. The POU domain is a bipartite DNA-binding structure. *Nature* **336**: 601–604.
- Szostak, J.W. 1992. In vitro genetics. *Trends Biochem. Sci.* **17**: 89–93.
- Thiesen, H.-J. and C. Bach. 1990. Target detection assay (TDA): A versatile procedure to determine DNA binding sites as demonstrated on SP1 protein. *Nucleic Acids Res.* **18**: 3203–3209.
- Treacy, M.N., X. He, and M.G. Rosenfeld. 1991. I-POU: A POU-domain protein that inhibits neuron-specific gene activation. *Nature* **350**: 577–584.
- Treacy, M.N., L.I. Neilson, E.E. Turner, X. He, and M.G. Rosenfeld. 1992. Twin of I-POU: A two amino acid difference in the I-POU homeodomain distinguishes as activator from an inhibitor of transcription. *Cell* **68**: 491–505.
- Treisman, J., P. Gönczy, M. Vashishtha, E. Harris, and C. Desplan. 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* **59**: 553–562.
- Treisman, J., E. Harris, and C. Desplan. 1991. The paired box encodes a second DNA-binding domain in the paired homeodomain protein. *Genes & Dev.* **5**: 594–604.
- Treisman, J., E. Harris, D. Wilson, and C. Desplan. 1992. The homeodomain: A new Face for the helix-turn-helix? *BioEssays* **14**: 145–150.
- Tuerk, C. and L. Gold. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**: 505–510.
- Umesono, K. and R.M. Evans. 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **57**: 1139–1146.
- Umesono, K., K.K. Murakami, C.C. Thompson, and R.M. Evans. 1991. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* **65**: 1255–1266.
- Verrijzer, C.P., M.J. Alkema, W.W. van Weperen, H.C. van Leeuwen, M.J.J. Strating, and P.C. van der Vliet. 1992. The DNA binding specificity of the bipartite POU domain and its subdomains. *EMBO J.* **11**: 4993–5003.
- Voss, J.W., L. Wilson, and M.G. Rosenfeld. 1991. POU-domain proteins Pit-1 and Oct-1 interact to form a heteromeric complex and can cooperate to induce expression of the prolactin promoter. *Genes & Dev.* **5**: 1309–1320.
- Wilson, T.E., R.E. Paulsen, K.A. Padgett, and J. Milbrandt. 1992. Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. *Science* **256**: 107–110.
- Wolberger, C., A.K. Vershon, B. Liu, A.D. Johnson, and C.O. Pabo. 1991. Crystal structure of a Mat $\alpha 2$  homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* **67**: 517–528.



## Cooperative dimerization of paired class homeo domains on DNA.

D Wilson, G Sheng, T Lecuit, et al.

*Genes Dev.* 1993, **7**:

Access the most recent version at doi:[10.1101/gad.7.11.2120](https://doi.org/10.1101/gad.7.11.2120)

---

### References

This article cites 81 articles, 25 of which can be accessed free at:  
<http://genesdev.cshlp.org/content/7/11/2120.full.html#ref-list-1>

### License

### Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

The advertisement features a dark blue background with a glowing DNA double helix on the left. The 'horizon' logo is in white, with 'a PerkinElmer company' in smaller text below it. To the right, the text 'Streamline your research with Horizon Discovery's ASO tool' is displayed in white.