A trans-acting peptide activates the yeast α1 represor by raising its DNA-binding affinity

Martha R. Stark1, Dominik Escher2 and Alexander D. Johnson1,3,4

1Department of Biochemistry and Biophysics, 2Department of Microbiology and Immunology, School of Medicine, University of California, San Francisco, CA 94143-0414, USA and 3Institute for Molecular Biology, University of Zurich, CH-8057 Zurich, Switzerland

© European Molecular Biology Organization 1621

The cooperative binding of gene regulatory proteins to DNA is a common feature of transcriptional control in both prokaryotes and eukaryotes. It is generally viewed as a simple energy coupling, through protein–protein interactions, of two or more DNA-binding proteins. In this paper, we show that the simple view does not account for the cooperative DNA binding of α1 and α2, two homeodomain proteins from budding yeast. Rather, we show through the use of chimeric proteins and synthetic peptides that, upon heterodimerization, α2 instructs α1 to bind DNA. This change is induced by contact with a peptide contributed by α2, and this contact converts α1 from a weak to a strong DNA-binding protein. This explains, in part, how high DNA-binding specificity is achieved only when the two gene regulatory proteins conjoin. We also provide evidence that features of the α1–α2 interaction can serve as a model for other examples of protein–protein interactions, including that between the herpes virus transcriptional activator VP16 and the mammalian homeodomain-containing protein Oct-l.

Keywords: homeodomain/protein–DNA interactions/protein–protein interactions/VP16/yeast cell-type determination

Introduction

In eukaryotes, proteins that regulate transcription typically act in specific combinations. A simple example of this principle is found in the specification of cell-types in the yeast Saccharomyces cerevisiae. In the a/α cell type, two homeodomain proteins, α1 and α2, form a heterodimer that binds with high affinity and specificity to a DNA sequence called the haploid-specific gene (hsg) operator (Goutte and Johnson, 1988, 1993, 1994; Dranginis, 1990). This operator is located upstream of many genes (collectively called the haploid-specific gene), and the binding of α1 and α2 to it recruits the SSN6–TUP1 complex, which represses transcription of each haploid-specific gene (Mukai et al., 1991; Keleher et al., 1992; Smith and Johnson, 1992; Komachi et al., 1994). These genes encode proteins required for the a and α cells to mate, and for regulators of a/α cell-specific functions (for reviews see Herskowitz et al., 1992; Johnson, 1995). Because α1 and α2 are present together only in the a/α cell type, the haploid-specific genes are derepressed in the other two yeast cell types, a and α cells.

The α1–α2 heterodimer has been well studied genetically, biochemically and structurally (e.g. Goutte and Johnson, 1993; Phillips et al., 1994; Li et al., 1995; Vershon et al., 1995). Some of these studies have utilized the minimal fragments of α1 and α2 necessary to heterodimerize and to bind tightly to DNA. The α2 minimal fragment includes the homeodomain (60 amino acid residues which fold into three α helices linked by two turns) and a short peptide tail (21 amino acid residues) which extends from the C-terminus of the homeodomain. The α1 minimal fragment includes only the 60-amino acid homeodomain. NMR studies have shown that the α2 tail is unstructured in the α2 monomer but folds into a short distorted α-helix upon contact with the α1 homeodomain (Phillips et al., 1991, 1994). An X-ray crystallographic study of the α1 and α2 minimal fragment heterodimer bound to DNA (Li et al., 1995) has revealed that this helix rests on top of the α1 homeodomain (that is, on the surface opposite that which contacts DNA), making a series of hydrophobic contacts. In the crystal structure both homeodomains contact the DNA which is strongly bent. The only contact between the α1 and α2 minimal fragments occurs through the α2 tail (Stark and Johnson, 1994; Li et al., 1995), although in the intact protein additional protein–protein contacts are made (Goutte and Johnson, 1993; Ho et al., 1994). Many additional genetic and biochemical experiments support the biological relevance of the crystal structure (reviewed by Andrews and Donoviel, 1995).

Despite these studies, it has been difficult to account for the high DNA-binding specificity of the heterodimer in terms of the individual DNA-binding properties of its two constituents. Under experimental conditions in which the specificity of the heterodimer for the hsg operator over non-specific DNA was at least 3000-fold, the α2 monomer exhibited a DNA-binding specificity of ~10-fold and the α1 monomer showed no reproducible sequence-specific DNA binding (Goutte and Johnson, 1993; Phillips et al., 1994).

In this paper, we address how the high DNA-binding specificity of the α1–α2 heterodimer is generated. In particular, we distinguish between two hypotheses: (i) heterodimerization is simply a coupling of the two monomers; and (ii) heterodimerization involves the instruction of one monomer by the other. Since no significant structural changes occur in the α2 homeodomain upon heterodimerization (Wolberger et al., 1991; Phillips et al., 1994; Li et al., 1995), the second hypothesis can be reduced to the more specific proposal that, upon heterodimerization, the α2 tail instructs the α1 homeodomain to bind specifically to DNA. Previous work (Mak...
and Johnson, 1993; Stark and Johnson, 1994) established the importance of the \( \alpha_2 \) tail for heterodimer formation, but did not distinguish between these two models.

To test this idea, we constructed a series of chimeric molecules consisting of the homeodomain of \( a1 \) linked covalently to the tail of \( \alpha_2 \). These two elements were joined by linkers designed to be of sufficient length and flexibility to permit an intramolecular interaction between the \( a1 \) homeodomain and the \( \alpha_2 \) tail. A prediction of the instructional hypothesis (but not of the simple coupling model) is that such chimeric \( a1 \) molecules should be able to bind tightly and specifically to DNA as monomers; we show that this is indeed the case. In a second set of experiments, we show that an \( \alpha_2 \) tail peptide supplied in solution can induce the \( a1 \) homeodomain to bind to DNA, a result that provides additional support for the instructional model. Finally, we provide evidence that features of the interaction between \( a1 \) and \( \alpha_2 \) also apply to other combinations of gene regulatory proteins.

Results

Construction of the \( a1::\alpha_2 \) chimeric proteins

The design of the chimeric \( a1 \) homeodomains (in particular the length of the linkers) was based on inspection of the X-ray crystal structure of the \( a1–\alpha_2 \) heterodimer bound to DNA (Li et al., 1995). They are shown schematically in Figure 1A. For two of the chimeric molecules, the \( \alpha_2 \) tail was attached via a linker to the C-terminus of the \( a1 \) homeodomain. In the X-ray structure the distance between the C-terminus of \( a1 \) and the N-terminus of the \( \alpha_2 \) tail is 32 Å. A linker of 11 amino acids (present in the \( a1::11::\alpha_2 \) chimera) should, in principle, span this distance if it is assumed that the linker is fully extended. A chimera with a linker of 16 amino acids (\( a1::16::\alpha_2 \)) was also constructed to accommodate some degree of folding in the linker. The linkers were composed of glycine and serine to provide both flexibility and solubility. A third chimeric molecule was constructed in which the \( \alpha_2 \) tail was attached via a linker to the N-terminus of the \( a1 \) homeodomain. The distance between the C-terminus of \( \alpha_2 \) and the N-terminus of the \( a1 \) homeodomain in the heterodimer crystal structure is only 13 Å, and a glycine/serine linker of 6 amino acids was used to span this distance.

\( a1::\alpha_2 \) chimeric proteins bind DNA as monomers

The four chimeric proteins depicted in Figure 1A were expressed in \( Escherichia coli \), purified to >90% homogeneity, and tested for their binding to a synthetic operator composed of two \( a1 \) half-sites (\( a1–a1 \) in Figure 1C). In contrast to the \( a1 \) homeodomain, which failed to bind DNA (Figure 2, lanes 5–7), all three chimeras exhibited efficient DNA binding in the 30–100 nM range (Figure 2, lanes 11–13, 17–19, 23–25). The chimeras \( a1::11::\alpha_2 \) (lanes 11–13) and \( a1::16::\alpha_2 \) (lanes 17–19) each formed two distinct protein–DNA complexes, whereas \( \alpha_2::6::a1 \) (lanes 23–25) produced only a single species. Based on a comparison with previous results (Smith and Johnson, 1992; Goutte and Johnson, 1993) and with the migration of the \( a1–\alpha_2 \) fragment heterodimer bound to DNA (Figure 2, lanes 2–4), we conclude that the \( \alpha_2::6::a1–DNA \) species and the faster migrating of the two \( a1::11::\alpha_2–DNA \) and \( a1::16::\alpha_2–DNA \) species are monomers of the \( a1 \) chimera bound specifically to DNA. Consistent with this assignment, DNease I footprinting of the chimeric proteins on an \( hsg \) operator (which contains an \( a1 \) half-site and an \( \alpha_2 \) half-site; see Goutte and Johnson, 1994) demonstrated that all three of the chimeric proteins bind only to the \( a1 \) half of the operator (Figure 3). In contrast, the \( a1–\alpha_2 \) fragment heterodimer binds to both the \( a1 \) and \( \alpha_2 \) sites of the operator (Figure 3).
this difference is that the placement of the linker on the α bound to DNA. As pointed out above, the binding of migrating species, which is very likely to be two monomers of a second molecule to the second a1 half-site, while the DNA binding of a1::

Fig. 2. Binding of the a1::α2 chimeras to the a1–a1 operator. The 32P-labeled DNA fragment (80 nucleotide pairs) containing the a1–a1 operator was incubated with the purified protein indicated for 30 min at room temperature and electrophoresed through a 5% native Tris-borate–EDTA polyacrylamide gel (as described in Stark and Johnson, 1994). Lane 1 contains labeled DNA alone. Lanes 2–4 contain 3 nM α2·hd + tail (α2·a128–210) in addition to a1·hd. The a1·hd alone (lanes 5–7) and the a1::α2 chimera homodimer (lanes 8–10) were included along with the a1–α2 fragment heterodimer in order to demonstrate the different mobility shifts expected for monomers and dimers bound to the DNA. a1::α2 (lanes 11–13) and a1::α2 (lanes 17–19) both give two shifts, consistent with monomeric and dimeric DNA binding, whereas α2::6::a1 (lanes 23–25) gives only one shift consistent with monomeric DNA binding. All three chimeras containing the α2 tail mutation have reduced DNA binding (lanes: 14–16, a1::11::α2H196S; 20–22, a1::16::α2H196S; 26–28, α2H196S:6::a1). The a1 homeodomain and chimera concentrations for each set of three reactions are 30, 100 and 300 nM.

Fig. 3. DNA binding of a1::α2 chimeras to an a1–α2 operator. DNase I protection of an 80 nucleotide pair 32P-labeled fragment containing the hsg operator. An operator with only one a1 binding site was chosen for this experiment so that monomeric binding of proteins would be revealed. Lane 1, 3μM a1·hd plus 100 nM α2a128–210; lane 2, 50 μM a1::α2; lane 3, no protein; lane 4, 50 μM α2·6::a1; lane 5, 100 μM a1::α2; lane 6, 50 μM a1::11::α2. The a1·α2 fragment heterodimer and the a1::α2 chimeric protein described previously as binding as a dimer (see text) were included as controls, to show protection of both half-sites in the operator. In contrast, α2::6::a1, a1::16::α2 and a1::11::α2 show protection only over the a1 half of the operator.

C-terminal placements permit the binding of a second monomer. Another possibility is that the dimeric species observed in this experiment arise, at least in part, from favorable contact between two monomers, and that this contact is suboptimal in a2::6::a1. In any case, the important point is that all three a1::α2 chimeras are capable of binding as monomers to specific DNA sites, whereas the a1 homeodomain itself shows no DNA binding under similar conditions.

We showed previously that a different a1::α2 chimera, one that contains the tail of α2 linked covalently to the C-terminal end of the a1 homeodomain but that lacks a linker, binds DNA only as a dimer (Stark and Johnson, 1994). For comparison, the behavior of this protein is shown in Figure 2 (lanes 8–10). According to the a1–α2–DNA crystal structure, this chimera (due to the absence of a linker) should not be capable of undergoing an intramolecular interaction to bring the α2 tail in contact with the proper surface of the a1 homeodomain. Thus, the only way for this chimera to efficiently bind the operator is through the interaction of the tail of one molecule with the homeodomain of a second molecule. We believe this is the explanation for the dimer requirement of this chimera. In contrast, the chimeras described in this paper are all capable of binding DNA as monomers. The predicted structures of the chimeras, based on the crystal structure of the a1–α2 fragment heterodimer, are shown in Figure 4A, B and C.

Specific point mutations in the tail reduce DNA binding by the a1::α2 chimeras

To rule out the possibility that the enhanced DNA-binding of the chimeric a1 proteins was due to non-specific contributions of the linker or of the tail, a set of additional chimeras was constructed, each of which contains a specific point mutation in the tail. Leu196 in the α2 tail is critical for a1–α2 function in vivo and in vitro (Strathern et al., 1988; Stark and Johnson, 1994). In the X-ray crystal structure of the a1–α2 heterodimer, this leucine makes an important contact with the a1 homeodomain (Li et al.,
M.R. Stark, D. Escher and A. D. Johnson

Fig. 4. Cartoons depicting interaction of the α2 tail with the α1 homeodomain. α1 is shown in red and α2 is shown in blue. The green ovals refer to the glycine/serine linkers (see Figure 1). (A) The α1–α2 heterodimer fragment crystal structure (Li et al., 1995). (B) Inferred structure of the two chimeras (α1::11::α2 and α1::16::α2) in which the α2 tail and linker were fused to the C-terminus of α1. (C) Inferred structure of the α2::6::α1 chimera, in which the α2 tail and linker were fused to the N-terminus of the α1 homeodomain. (D) Inferred structure of the α1 homeodomain bound by wild-type α2 tail peptide. For each of these cases, we have provided evidence that the interaction of the α2 tail with the homeodomain of α1 induces a conformational change, which stimulates the binding of the α1 homeodomain to hsg operator DNA. The contacts between α1 and DNA occur through helix 3 and the loop between helices 1 and 2 (Li et al., 1995). As described in the text, it has been proposed that the conformational change in α1 occurs in this loop.

1995). In all three chimeras, this leucine was changed to serine, and in all three cases (Figure 2, lanes 14–16, a1::11::α2L196S; lanes 20–22, a1::16::α2L196S; and lanes 26–28, α2L196S::6::a1) DNA binding was significantly reduced compared with chimeras that carried the wild-type tail. This result demonstrates that efficient DNA binding by the α1 homeodomain chimeras specifically requires a functional α2 tail.

The α2::6::α1 chimeric protein shows sequence-specific DNA binding
The DNase protection experiments of Figure 3 showed that the three α1 chimeras specifically recognized the α1 half-site. We verified the sequence-specific DNA-binding of one chimera, α2::6::α1, using an additional technique. Using a gel mobility assay, we directly compared its affinity for a DNA fragment that contained the synthetic α1–α1 binding site with its affinity for the same fragment lacking these sites (Figure 5). Both α2::6::α1 (lanes 9–11) and the α1–α2 fragment heterodimer, used as a control (lanes 3–5), bind specifically to the α1–α1 fragment, but not to the operator that lacks α1 sites (lanes 6–8 and 12–14). These results verify the footprinting results, showing that the α2::6::α1 chimera has a marked preference for a known α1 binding site over other DNA.

An α2 tail peptide is sufficient to induce the α1 homeodomain to bind DNA
It seemed plausible, based on the results described above, that the isolated tail of α2 could induce the α1 homeo-

Fig. 5. DNA-binding specificity of the α2::6::α1 chimera. For this electrophoretic mobility shift experiment, an 80 nucleotide pair 32P-labeled fragment containing the α1–α1 operator and the same labeled fragment with the operator deleted (consisting of 51 nucleotide pairs) were utilized. The two operators have different mobilities in the gel due to their size difference. Lanes 1, 3–5 and 9–11 contain the α1–α1 fragment while 2, 6–8 and 12–14 contain the fragment that lacks specific α1-binding sites. Lanes 1 and 2 lack protein and the other lanes have the indicated protein or proteins added. The binding conditions and protein concentrations were the same as those used in Figure 2.
 Activation of DNA-binding in a homeodomain protein

Fig. 6. A 19 amino acid α2 tail peptide, supplied in trans, induces α1 to bind the α1–α1 operator. Each lane contains 30 nM α1 hd, which does not bind DNA on its own (lane 1). In addition to α1 hd, lanes 2–6 contain successive 2-fold increases of the wild-type α2 tail peptide beginning with a concentration of 0.15 mM in lane 2 and ending with 2.5 mM in lane 6. Both mutant peptides, present in the same concentrations as the wild-type peptide, are reduced in their ability to induce α1 DNA binding (L196S, lanes 7–11; L196A, lanes 12–16). The binding conditions are the same as those of Figure 2, except that the incubations of DNA and protein were carried out at 4°C, as was the electrophoresis.

Fig. 7. Sequence alignment of the ‘tails’ experimentally attached to the α2 homeodomain. Amino acid sequence of the S.cerevisiae (S.c.) and K.lactis (K.l.) α2 C-terminal tails correspond to the C-terminal 21 and 12 residues, respectively, of the two proteins. The VP16 sequence numbers correspond to residues 371–389 in the full-length VP16 protein (1–490). The residues of the S.cerevisiae α2 tail shown in bold are those that interact with the surface of the α1 hd. The corresponding residues in the other two tails are also shown in bold for comparison.

The results illustrated in Figure 6 show that the amount of DNA bound by the homeodomain increases as the concentration of wild-type peptide is raised (lanes 2–6), with half-maximal stimulation reached at a peptide concentration of ~0.3 mM. This value is in excellent agreement with the $K_D$ of 0.2–0.3 mM seen for the interaction of the α2 and α1 fragments as measured by NMR spectroscopy (Baxter et al., 1994; Phillips et al., 1994). The two mutant peptides also stimulate α1 binding (Figure 6, lanes 7–11 and 12–16), but to a significantly lesser extent. In other experiments, the difference between the wild-type and mutant peptides was less pronounced, suggesting that the α1–peptide interaction is highly sensitive to the conditions employed. The predicted structure of the α1 homeodomain–peptide complex, based on the crystal structure of the α1–α2 heterodimer fragment, is shown in Figure 4D.

Can heterologous ‘tails’ also mediate cooperative binding of α1 and α2 to DNA?

We have presented evidence that the α2 tail, in addition to joining α1 and α2 together to form the heterodimer, acts as a ligand to convert the α1 homeodomain to a high-affinity DNA-binding form. In this section, we test the generality of the tail–homeodomain interaction by testing whether a heterologous tail can substitute for the α2 tail in bringing about the cooperative binding of α1 and α2. We tested two heterologous tails for this activity. The first was from the α2 protein of Kluyveromyces lactis, a yeast closely related to S.cerevisiae. As judged by sequence comparison, the K.lactis α2 tail is slightly shorter than that of S.cerevisiae (Figure 7), but the last several residues of the S.cerevisiae α2 tail are dispensable for its function (see above). The four hydrophobic residues that form the basis of the interaction between the S.cerevisiae α2 tail and the α1 homeodomain are also hydrophobic in the K.lactis α2 tail, but only one of the 12 tail positions (Leu200 in α2) contains the same amino acid in both proteins.

The second tail we tested derives from the herpes simplex virus transcriptional regulator, VP16. Although not itself a homeodomain protein, VP16 interacts with the mammalian Oct-1 homeodomain to maximally activate transcription of some of its target genes. Baxter et al. (1994) and Li et al. (1995) have pointed out that the region of VP16 shown to interact with Oct-1 (Werstuck and Capone, 1989a,b; Stern and Herr, 1991; Walker et al., 1994; Shaw et al., 1995; Lai and Herr, 1997) bears rough amino acid similarity to that of the α2 tail (Figure 7). Moreover, the results of mutagenesis experiments (Lai et al., 1992; Pomerantz et al., 1992) have suggested that the surface of Oct-1 contacted by VP16 is approximately equivalent to the surface of α1 contacted by α2.

Analysis of the K.lactis α2 tail

We substituted the C-terminal tail of the α2 protein from S.cerevisiae with that of the K.lactis α2 to give an α2 hd::K.l. tail chimera. We expressed this chimeric protein in E.coli and showed that it was capable of binding DNA cooperatively with the S.cerevisiae α1 homeodomain protein (Figure 8A). In this DNA-binding experiment, the α1 homeodomain on its own shows no detectable DNA binding, but in the presence of the α2 hd::K.l. tail chimeric protein, significant DNA binding is observed. The migration of the complex is consistent with a heterodimer
bound to DNA. Approximately 10-fold more α2 hd:K.lactis tail protein than S.cerevisiae α2 homeodomain protein is required to bind an equivalent amount of DNA in the presence of the same concentration of α1 (data not shown), indicating that the interaction between the K.lactis α2 tail and the S.cerevisiae α1 homeodomain is less favorable than that between the two S.cerevisiae proteins.

To determine whether the K.lactis α2 tail was interacting with the α1 homeodomain in a manner analogous to that of the S.cerevisiae α2 tail, we changed one of the hydrophobic residues in the K.lactis tail, isoleucine 218, to serine. In S.cerevisiae, the equivalent mutation (L196S) disrupts the interaction between α1 and α2 (see above). The chimera α2 hd::K.l.lactis fails to bind cooperatively with the α1 homeodomain (Figure 8A). From these results, we conclude that the C-terminal tail of K.lactis α2 can at least partially substitute for the S.cerevisiae α2 tail. Thus, the interaction with the α1 homeodomain is maintained even though the two tails are identical at only a single position.

Analysis of the VP16 ‘tail’

To test further the generality of the α1 homeodomain–α2 tail interaction model, we grafted a region of the herpes virus activator VP16, onto the α2 homeodomain, in place of α2’s own tail (α2 hd::VP16). This region of VP16 is predicted to form an amphipathic helix and to interact with the exposed surface of the Oct-1 homeodomain protein (Hayes and O’Hare, 1993; Lai and Herr, 1997). In the VP16 tail, four hydrophobic residues correspond to the hydrophobic residues in the α2 tail that form the surface with which it interacts with the α1 homeodomain; only two of these four residues are identical between the two proteins. Overall, this region of VP16 is identical to the α2 tail at these two positions out of a total length of 19.

The DNA-binding experiment of Figure 8B shows that this α2 hd::VP16 chimera is also capable of binding to DNA cooperatively with the α1 homeodomain. However, the binding is ~50-fold weaker than that observed for the wild-type α2 protein. As discussed above, the VP16 and α2 tails differ considerably in amino acid sequence, and this difference may account for the poorer interaction seen with this chimera when compared with the α2 hd::K.l.lactis tail.

When the residue in VP16 that corresponds to Leu196 in the α2 tail is mutated to serine (I377S), the interaction of the mutant chimeric protein with the α1 homeodomain is reduced by more than a factor of 10 (Figure 8B), indicating that this residue plays a crucial role in the interaction with α1 and suggesting that the VP16 tail interacts with the α1 homeodomain in a manner similar to that of the α2 tail.

Discussion

α1 and α2 are homeodomain proteins that regulate cell identity in the budding yeast S.cerevisiae. Like other proteins that bind DNA through homeodomains, α1 and α2 each lack the specificity and affinity to select target DNA efficiently on their own. However, they form a heterodimer in solution which can then bind tightly and specifically to the hsg operators, thereby turning off transcription of the haploid-specific genes (for review see Johnson, 1995). In the DNA-bound state, the homeodomains of α1 and α2 both make extensive contacts with the hsg operator (Li et al., 1995).

The work presented in this paper shows that heterodimer formation is not simply the joining together of the α1 and α2 homeodomains. Heterodimerization also induces changes in each monomer, and the change in α1 is crucial for the DNA binding specificity of the heterodimer. A key contact in the heterodimer is formed by a flexible tail that extends from the α2 homeodomain and that becomes ordered upon contact with an exposed surface of α1 (Phillips et al., 1994; Li et al., 1995). This conformational change in the α2 tail has been well studied by NMR and X-ray crystallographic methods. In the α2 monomer, this tail is unstructured, but it assumes a distorted α-helix upon contact with α1 (Phillips et al., 1991, 1994; Wolberger et al., 1991; Li et al., 1995). However, this change in α2 has no apparent consequence on the DNA-binding properties of α2 (Mak and Johnson, 1993), and the overall structure of the α2 homeodomain (excluding the tail) is the same whether or not it is complexed with α1.

In this paper, we have provided evidence that contact with the α2 tail induces a change in the α1 homeodomain which renders it competent to bind tightly and specifically to DNA; without this instructional change, the α1 homeodomain binds DNA only weakly, if at all. We propose that the change in the DNA-binding properties of α1 induced by contact with α2 results from a conformational change in α1. What might this conformational change be? Baxter et al. (1994) showed by applying isotope-edited NMR spectroscopy to 15N-labeled α1 homeodomain and 14N-labeled α2 homeodomain plus tail that the resonances of many positions in the α1 homeodomain changed upon addition of α2. Some of these changes can be accounted for by direct contact by the α2 tail; however, others lie in positions more distant from the sites of direct contact. A cluster of α2-induced changes lies in the loop between helix 1 and helix 2 of the α1 homeodomain. As seen in the heterodimer crystal structure, this loop makes contact with both the DNA (via a water molecule) and with the α2 tail, and it is plausible that contact with the α2 tail repositions this loop to maximize the affinity of α1 for DNA (Li et al., 1995). A network of protein–DNA contacts involving the backbone of this loop is clearly observed in the heterodimer–DNA crystal structure, and even a subtle change in the conformation of this loop (predicted to occur upon dissociation of the α2 tail) could disrupt this network of contacts and significantly weaken the affinity of α1 for DNA.

We have also provided evidence that the principles underlying the α1–α2 interaction apply to the association of other homeodomain proteins. In particular, we have shown that two other tails can substitute for that of α2 in promoting heterodimer formation and DNA binding of α1 and α2. All three tails used in this study (from S.cerevisiae α2, from K.lactis α2 and from herpes virus VP16) are predicted (or in the case of S.cerevisiae α2, known) to form amphipathic helices. The VP16 region contains only two residue positions identical to those in S.cerevisiae α2 and the K.lactis tail only a single identical residue. However, the overall conservation of residue type across the entire tail region is higher in the K.lactis tail than in VP16, and this difference may explain why, of the two,
the \textit{K.lactis} tail is more efficient at cooperating with \textalpha. Despite its lower efficiency, we believe that this short region of VP16 is able to interact with the \textalpha homeodomain in a manner similar to that of the \textalpha2 tail. This idea is based in part on the observation that a specific point mutation in the VP16 tail, located at a position corresponding to one crucial for the \textalpha1–\textalpha2 interaction, destroys the cooperative binding with \textalpha. Two possibilities, not mutually exclusive, could account for the observation that the heterologous tails do not function as well in mediating \textalpha1–\textalpha2 cooperative DNA binding as does the \textit{S.cerevisiae} \textalpha2 tail. First, the affinity between the heterologous tail and the surface of the \textalpha homeodomain may differ, with the \textit{S.cerevisiae} \textalpha2 tail having the highest affinity. Secondly, the appropriate instructional change in \textalpha1 may be induced only by the \textalpha2 tail; the two other tails may effectively link \textalpha1 and \textalpha2 together, but may not fully induce the correct change in \textalpha1.

The fact that the VP16 tail can interact with the \textalpha homeodomain strongly supports the view that this region of VP16 interacts with the Oct-1 homeodomain in a similar manner. A number of additional observations also support this view. Mutations in this region of VP16 disrupt cooperative binding to the DNA with Oct-1 without affecting the ability of VP16 to interact with another factor, HCF (Werstuck and Capone, 1989a,b; Stern and Herr, 1991; Walker \etal, 1994; Shaw \etal, 1995; Lai and Herr, 1997). Peptides corresponding to this region of VP16 can affect the DNA-binding specificity of the Oct-1 homeodomain, as does the intact VP16 protein (Stern and Herr, 1991). Similar peptides can inhibit formation of the Oct-1–VP16 complex, presumably by binding to the same surface of Oct-1 as is normally bound by the full-length VP16 protein (Haigh \etal, 1990; Hayes and O’Hare, 1993; Wu \etal, 1994). Moreover, the surface of the Oct-1 homeodomain that contacts VP16 has been mapped through mutagenesis studies, and it corresponds to the same region of the \textalpha1 homeodomain that is contacted by \textalpha2 (Figure 9; Lai \etal, 1992; Pomerantz \etal, 1992).

The observations presented in this paper support the hypothesis that the principles underlying the \textalpha1–\textalpha2 interaction are conserved among other homeodomain proteins. There are now numerous examples of cooperative interactions involving homeodomain proteins, some of which might also be mediated by these same types of interactions, specifically by an amphipathic helix binding to the surface of its partner homeodomain. Only a limited number of solvent-exposed residues are available on a homeodomain bound to DNA, so it is plausible that the interaction between \textalpha1 and \textalpha2 (an \alpha-helix resting on helices 1 and 2) may have appeared early during the evolution and diversification of homeodomain proteins.

There are also indications that additional pairs of homeodomain proteins undergo conformational changes upon heterodimerization. Chan and co-workers (1996) have proposed that the YPWM hexapeptide of the fly labial protein, a homolog of the mouse Hoxb1 protein, inhibits its DNA binding and that interaction with Exd removes this inhibition. A second example is found in the mammalian Pbx1 protein. The affinity of the Pbx1 protein is enhanced by YPWM-containing peptides derived from several of the partners of Pbx1, and one of several models that could explain this result is a peptide-induced conformational change in the Pbx1 protein (Knoepfler and Kamps, 1995; Peltenburg and Murre, 1996; Sanchez \etal, 1997). Conformational changes induced by the heterodimerization of Hoxb8 and Pbx1 have been determined by circular dichroism spectroscopy (Sanchez \etal, 1997).
1997), although the exact nature of these changes and their consequences for DNA binding remain to be determined.

Finally, the idea that α2 carries a ligand that increases the affinity of α1 for DNA is similar in principle to the many cases of small molecules that directly activate the binding of proteins to DNA. Examples include such ligands as cAMP for E.coli CAP (Beckwith, 1987; Reznikoff, 1992; Ebright, 1993) and tryptophan for the E.coli trp repressor (Yano sky and Crawford, 1987; Somerville, 1992). In the case of α1, the small molecule is, in a sense, carried by the partner protein α2. A ligand-induced change seems an efficient way of ensuring that α1 is inactive in cells that lack α2 (α cells), but becomes activated only when α2 is also present, the condition that determines the α1 α2 cell-type.

Materials and methods

Construction of expression plasmids

The DNA encoding the glycine/serine linkers was synthesized as complementary oligonucleotides, annealed, and ligated between DNA encoding the α1 hd and DNA encoding the tail of α2 (pMSS5; see Stark and Johnson, 1994). The linkers are composed of alternating (glycine)2 and (serine)2. The oligonucleotides that made up the linkers in the α1::α2 chimeras are as follows: α1::16::α2–GATCTAAAGTGTTGCTCTTGGCTGGCCTCTCCCTGGGCTC. The first two oligonucleotides have an overhanging GATC at the 5′ end of each oligonucleotide of the pair, and were cloned into the BglII site at the junction of the α1 homeodomain and the α2 tail, yielding pMS74 and pMS87. The third linker oligonucleotide was part of a larger oligonucleotide which contains the α2 tail sequence (aa 189–210) immediately upstream of the linker. This oligonucleotide pair has an overhanging GATC at the 5′ end of the α1 homeodomain (pMS4; see Philipps et al., 1994), resulting in pMS76. The L196S mutants of the α1::11::α2 and α1::16::α2 chimeras were made by inserting the original oligos into the BglII site at the junction of the α1 homeodomain and the α2 tail (pMS18), resulting in pMS91 and pMS92. The α2L196S::α1 was generated by incorporating the codon change into the oligonucleotide pair and cloning into pMSS4 (pMSS7).

All α2 chimeric constructs were made by replacing the wild-type α2 tail with an oligo duplex consisting of either the K.lactis α2 tail (nucleotides corresponding to residues 212–223, MRS124 5′-TCGACGAAAAGAAAAAACAAACAATTACGGGTC-3′), or the K.lactis α2 tail (pMS22). The tail oligo duplexes contain a SalI overhang on the 5′ end and a BamHI overhang on the 3′ end, with the first 19 nucleotides corresponding to α2 sequence upstream of the tail.

A tail site was introduced into α2 19 bp upstream of the tail by site-directed mutagenesis (pMS20). The SalI site was removed from the yeast CEN ARS vector pAV115 containing the MATα locus (pMS21). pMS21 was cut with BglII and BamHI to remove the majority of the α2 gene. This was replaced with the corresponding BglII–BamHI fragment from pMS20 that contains α2 with the SalI site upstream of the α2 tail (pMS22). The SalI–BamHI fragment containing the α2 tail was removed from pMS22 and replaced with either the wild-type or mutant Kl. lactis α2 tail or VP16 tail (pMS104, pMS105, pMS85, pMS103, respectively). The α2 and VP16 versions of these chimeras were generated by PCR using pMS104, 105, 85 and 103 as templates. The 5′ primer for all of these PCR reactions introduces an NdeI site at the beginning of the α2 hd (5′-GATAACAACATATGAAAACCTACAGAG). The 3′ primers for all these PCR reactions are specific for each tail: Kl. lactis α2 tail–MRS128 5′-CGGCCGATCTTAATAGGTTTCTCCTAGGGGATCGTCCGGG-3′, Kl. lactis α2 tail–MRS128 5′-CGGCCGATCTTAATAGGTTTCTCCTAGGGGATCGTCCGGG-3′. The resulting PCR fragments were cut with NdeI and BamHI and cloned into these sites in the bacterial expression vector, pHS40, under the control of the T7 promoter (Studier and Moffat, 1986), resulting in pMS109–α2 hd::Kl. tail, pMS110–α2 hd::Kl. tail, pMS89–α2 hd::VP16, pMS108–α2 hd::VP16.775.

Peptides

α2 tail peptides were synthesized by California Peptide Research, Inc. All three peptides are identical except for position 196; they are 19 amino acids in length, beginning at residue 189 of α2 and ending at residue 207. The amino acid sequence of the wild-type α2 tail peptide is TITIAPELADILLGSQPLAK. Residue 196 is in bold. Peptides were HPLC-purified, resuspended in H2O, and concentrations were determined by the quantitative Ninhydrin assay (Sarin et al., 1981).

Operators

The α1–α2 operator has the same spacing and binding site orientation as a naturally occurring hsg operator (which contains an α1 and α2 binding site), except the α2 binding site of the hsg operator has been replaced by a second α1 binding site (Goutte and Johnson, 1994). The two α1 binding sites are separated by six base pairs and, for the experiments shown in Figures 2, 5 and 6, are contained within an 80-bp DNA fragment. A second DNA fragment, identical in sequence to the α1–α2 operator-containing fragment except that it contains no specific α1-binding sites, was used in the experiment of Figure 5. The removal of the α1-binding sites results in a 51 nucleotide pair DNA fragment.

Protein purification

The α1 hd protein and all α1:α2 chimeric proteins were overexpressed in E.coli strain BL21(DE3)pLysS. Protein purification from cell lysates was by adhesion to a cation-exchange resin (Sephadex SP-C50, Pharmacia) followed by elution with a NaCl gradient (Phillips et al., 1994). The α2 hd α2 tail fragment was a gift from A.Veshon.

All α2 hd chimeric proteins were present in bacterial extracts. These extracts were made from the protease-deficient E.coli strain CAG25D7 7 ml/g lysis buffer {100 mM Tris–HCl, pH 8.0, 1 mM EDTA, 10 mM β-mercaptoethanol, 500 mM NaCl and 0.1 mM [4-(2-aminoethyl)benzenesulfonyl fluoride] (AEBSF) (Calbiochem)}, sonicated to lyse the cells and centrifuged at 30 000 × g for 40 min. The supernatant was used in the gel shift assays after quantitating the amount of hsd 2 hd protein in each extract by SDS gel followed by Coomassie Blue staining.

DNA-binding assays

For the electrophoretic mobility shifts containing only purified proteins (α1:α2 chimeric proteins), proteins were incubated with a 32P-labeled DNA fragment for 30 min at room temperature and electrophoresed through a 5% native Tri–borate–EDTA polyacrylamide gel as described previously (Stark and Johnson, 1994). For the electrophoretic mobility shifts containing overexpressed α2 hd chimeric proteins in bacterial extracts, the extracts were incubated, either in the presence or absence of purified α1 hd protein, with labeled DNA for 45 min on ice before electrophoresis. The binding conditions for the α2 tail peptides and labeled α1–α2 operator were the same as in the other α1::α2 experiments except that the incubations of DNA and protein were carried out at 4°C, as was the electrophoresis.

The DNase I protection experiment was carried out under the same conditions as the mobility shift assays using purified proteins, except that 50–100 times more DNA was used (a 5′-labeled DNA fragment containing an hsg operator), and the binding buffer contained no glycerol or E.coli genomic DNA, but was supplemented with 10 mM CaCl2, and 2.5 mg/ml calf thymus DNA. Reactions were cleaved for 10 min at room temperature with 1.5 mg DNase I ( Worthington) and then stopped and precipitated with 1.6 M ammonium acetate. Samples were electrophoresed through a 10% denaturing TBE gel.
Activation of DNA-binding in a homeodomain protein

Acknowledgements
We thank T.Li and C.Wolberger for providing us with the coordinates for the α2-DNA crystal structure prior to publication, S.Rader, S.Triezenberg and members of the Johnson laboratory for helpful discussions and comments on the manuscript, S.Rader for help with the structure representations and W.Schaffer for his support of D.E. This work was funded by a NIH Grant RO1 GM37049 to A.D.J.

References
Received November 2, 1998; revised and accepted January 27, 1999