Crystal Structure of the MATa1/MATα2 Homeodomain Heterodimer Bound to DNA

Thomas Li, Martha R. Stark, Alexander D. Johnson, Cynthia Wolberger*

The *Saccharomyces cerevisiae* MAT**a**1 and MATα2 homeodomain proteins, which play a role in determining yeast cell type, form a heterodimer that binds DNA and represses transcription in a cell type–specific manner. Whereas the α2 and **a**1 proteins on their own have only modest affinity for DNA, the **a**1/α2 heterodimer binds DNA with high specificity and affinity. The three-dimensional crystal structure of the **a**1/α2 homeodomain heterodimer bound to DNA was determined at a resolution of 2.5 Å. The **a**1 and α2 homeodomains bind in a head-to-tail orientation, with heterodimer contacts mediated by a 16-residue tail located carboxyl-terminal to the α2 homeodomain. This tail becomes ordered in the presence of **a**1, part of it forming a short amphipathic helix that packs against the **a**1 homeodomain between helices 1 and 2. A pronounced 60° bend is induced in the DNA, which makes possible protein-protein and protein-DNA contacts that could not take place in a straight DNA fragment. Complex formation mediated by flexible protein-recognition peptides attached to stably folded DNA binding domains may prove to be a general feature of the architecture of other classes of eukaryotic transcriptional regulators.

Homeodomain proteins constitute a superfamily of DNA binding proteins that play critical roles in gene regulation and development in many eukaryotic species. These proteins have in common a conserved 60–amino acid DNA binding domain that has been well characterized structurally, biochemically, and genetically (*1*, *2*), whereas other portions of the proteins are quite divergent. The fold adopted by the homeodomain, as uncovered in structural studies of *Drosophila* (*3*–*5*), yeast (*6*, *7*), and human (*8*) homeodomains, consists of three α helices and an NH_2-terminal arm. Homeodomains bind DNA by inserting the third of these three α helices into the major groove of the DNA, while the NH_2-terminal arm contacts bases in the adjacent minor groove. Biochemical studies have shown that isolated homeodomains, which typically bind DNA as monomers, often exhibit only a relatively modest degree of DNA sequence selectivity (*9*–*11*). In the case of the yeast α2 homeodomain protein, we know that the specificity and affinity with which α2 binds DNA is augmented by its association with either of two partner proteins: the product of the MAT**a** locus, **a**1, or the non–cell type–specific protein MCM1. The **a**1/α2 and α2/MCM1 complexes each bind to a distinct set of DNA sites in the yeast genome, causing repression of the adjacent genes. Combinatorial control by **a**1 and α2 provides a means to achieve cell type–specific regulation of a large set of genes since **a**1 and α2 are expressed together in only the **a**/α diploid yeast cell type (*12*).

One of the partner proteins of α2, **a**1, is also a member of the homeodomain superfamily (*13*). In the diploid **a**/α cell type, these two homeodomain proteins form a heterodimer that binds to sites upstream of haploid-specific genes (hsg) (*14*–*16*). In the absence of α2, the **a**1 protein exhibits no detectable specific binding to DNA (*11*). However, the presence of **a**1 in solution dramatically raises the affinity of α2 for hsg operators. The cooperative binding of α2 with **a**1 depends on the 21-residue COOH-terminal tail of α2, which is located immediately COOH-terminal to its homeodomain (*17*, *18*). Deletion of the COOH-terminal tail renders α2 incapable of cooperative binding with **a**1 in vitro and of repressing the haploid-specific genes in vivo (*18*). This tail is required for interaction with the **a**1 partner protein only, as its absence does not affect the cooperative interaction of α2 with MCM1 (*18*). The interaction of the tail of α2 with the **a**1 protein is quite specific; for example, the tail fails to interact with α2 itself, a homeodomain closely related to **a**1 (*17*). Splicing this peptide onto the *Drosophila* engrailed homeodomain renders engrailed capable of cooperative interaction with **a**1 (*17*). Additional interactions between **a**1 and α2 are mediated by the NH_2-terminal domains of the respective proteins, which have been proposed to contact one another by way of a coiled-coil interaction (*19*). Deletion analysis of **a**1 and α2 has shown that their cooperative binding to the hsg operator requires only the homeodomain of **a**1 and the homeodomain plus the tail of α2 (*20*). These fragments, which were used for

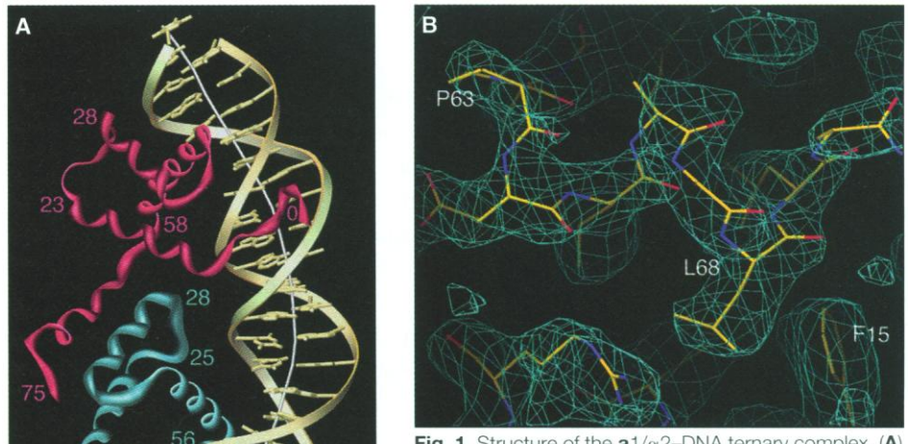


Fig. 1. Structure of the **a**1/α2–DNA ternary complex. (**A**) Overall model of the complex. The α2 protein is shown in red and **a**1 is shown in blue. Sequence numbers for the **a**1 and α2 proteins follow the same color scheme. The white line indicates the local DNA axis. This figure, as well as Figs. 3B, 4, A and C, 6, B to D, and 7 were generated with the program SETOR (*59*). (**B**) Simulated annealing omit map showing the COOH-terminal tail of α2 bound to **a**1. Residues 59 to 74 of α2 were deleted and the model was subjected to limited simulated annealing refinement in order to remove phase bias. The electron density map shown was calculated with $2F_o - F_c$ coefficients and phases from the resulting model and is contoured at 1.0 σ.

T. Li is in the Department of Biophysics and Biophysical Chemistry and C. Wolberger is in the Department of Biophysics and Biophysical Chemistry and the Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205–2185, USA. M. R. Stark and A. D. Johnson are in the Department of Microbiology and Immunology, University of California School of Medicine, San Francisco, CA 94143–0502, USA.

*To whom correspondence should be addressed.

the structural studies described here, bind DNA with 10-fold reduced affinity as compared with the full-length proteins (*11*, *20*). We refer to the complex formed by these fragments as the **a**1/α2 heterodimer. The equilibrium constant for dissociation of this **a**1/α2 heterodimer into monomers is 2×10^{-4} M (*20*).

We previously described the x-ray crystal structure of the α2 homeodomain bound to DNA (*6*). In the absence of its partner proteins, the α2 homeodomain binds as a monomer to nearly straight B-form DNA. Although the fragment of α2 crystallized contained the COOH-terminal tail that mediates heterodimerization with **a**1, the 21 residues of the tail were disordered in the structure of the α2-DNA complex. Solution nuclear magnetic resonance (NMR) studies show that the COOH-terminal tail is unstructured in solution as well, becoming ordered only upon heterodimerization with **a**1, when part of it forms a short α helix (*7*, *20*).

In order to uncover how a flexible peptide tail mediates specific heterodimer formation between two homeodomains, we determined the three-dimensional structure of a ternary complex containing an **a**1/α2 heterodimer bound to DNA. This crystal structure, determined at a resolution of 2.5 Å, shows that the **a**1 and α2 homeodomains bind in tandem to a 21–base pair (bp) DNA fragment. In the ternary complex, 16 residues of the COOH-terminal tail of α2 undergo a conformational change, becoming ordered and contacting the **a**1 homeodomain at a surface that does not participate in DNA binding. The heterodimer induces a pronounced DNA bend that is required for contacts between the two proteins. The manner in which **a**1 and α2 heterodimerize represents a possible mechanism by which other transcriptional regulators can associate with one another on the DNA.

Overview of the ternary complex. The **a**1/α2–DNA complex (Fig. 1A) contains the **a**1 and α2 proteins bound to a 21-bp fragment of bent duplex DNA. The structure of the complex was determined as described (Table 1); a representative view showing the fit of the model to the electron density map is shown in Fig. 1B. The sequence of the DNA site was derived from the sequences of 14 in vivo **a**1/α2 binding sites (Fig. 2C) and differs in the four central base pairs from the consensus sequence. Of a number of different DNA sequences and oligonucleotide lengths, this DNA site yielded the best crystals. The sequences of the proteins and the DNA, and the numbering scheme used to describe them, are shown in Fig. 2. The **a**1 and α2 homeodomains bind in a tandem orientation to one face of the DNA, burying 2300 Å^2 of protein and DNA surface area. The tandem binding of the two homeodomains, which had been predicted on the basis of chemical protection experiments (*16*), is in agreement with the observed pattern of DNA protection from hydroxyl radical attack and methylation in the presence of bound **a**1/α2 (Fig. 2D).

The α2 protein contacts the **a**1 homeodomain with a peptide tail located COOH-terminal to the α2 homeodomain. As compared with the structure of α2 alone bound to DNA, an additional 16 residues are ordered in the ternary complex. This COOH-terminal tail, consisting of residues 59 to 74, extends from the end of helix 3 of the α2 homeodomain (Fig. 1A). Residues 59 to 62 contain a short stretch of extended chain, followed by two turns of an amphipathic helix (residues 63 to 69), which contacts the **a**1 homeodomain on the face opposite to that which binds DNA. The helix in the tail of α2 packs between helices 1 and 2 of **a**1, with the axes of all three helices roughly parallel. The α2 tail helix is somewhat distorted, with deviations from ideal hydrogen bonding geometry at residues 64 and 65. As predicted in biochemical and NMR studies (*20*), all contacts with the **a**1 homeodomain are mediated by the COOH-terminal extension of α2; the homeodomain of α2 (residues 0 to 59) forms no direct contacts with the **a**1 protein.

The model of the ternary complex contains residues 0 to 74 of α2, including the

Table 1. Crystallographic analysis. The purification and crystallization of the **a**1/α2–DNA complex and the method of flash-freezing the crystals have been described (*51*). The complex crystallizes in space group $P6_1$ with unit cell dimensions $a = b = 132.8$ Å and $c = 45.71$ Å. The crystals form with one complex per asymmetric unit and contain 68 percent solvent. Diffraction from these crystals extends to a maximum resolution of 2.4 Å, with diffraction along the *c* axis falling off in intensity beyond 2.7 Å. Heavy atom derivatives were prepared by substituting 5-iodouracil for thymine at base 1, 11, or 22 of the DNA (IdU^1, IdU^{11}, and IdU^{22}). X-ray diffraction data collection, processing, and multiple isomorphorous replacement (MIR) phase calculation and refinement were as described (*52*). The MIR phases were used to calculate an electron density map at 2.8 Å resolution, which was improved by one round of solvent flattening (*53*). A nearly complete model of the complex was fit to this map by means of the O graphics program (*54*). A statistically random selection of 10 percent of the total reflection data was excluded from the refinement and used to calculate the free *R* factor (R_{free}) as a monitor of model bias (*55*). The model was subjected to several rounds of positional and simulated-annealing refinement with X-PLOR (*56*). Several missing side chains and residues were added following inspection of $2F_o - F_c$ and $F_o - F_c$ maps and a phase-combined map calculated with SIGMAA (*57*). The analysis was continued with 120 cycles of positional and constrained temperature factor refinement, yielding an *R* factor at 2.8 Å resolution of 22.4 percent and an R_{free} of 29.2 percent. The model was then further refined at 2.5 Å resolution against the derivative data set, IdU^{11}. In the resolution range from 2.8 to 2.5 Å, the IdU^{11} data set contained 94 percent of the expected data; of those reflections recorded, 81 percent of the intensities were greater than 2σ. After rigid body, positional, and simulated annealing refinement, an additional five residues in the COOH-terminal tail of α2 were fit to $2F_o - F_c$ and $F_o - F_c$ maps. Simulated annealing omit maps (*58*) were calculated at many stages of the refinement to verify the placement of residues in the electron density map. Water molecules were included at the final stage of refinement, based on the presence of peaks in difference electron density maps of at least 3σ in significance. All water molecules have *B* factors of less than 50 Å^2 and participate in at least one hydrogen-bonding interaction. The model of the complex presented here contains 1851 atoms and 58 water molecules. The average atomic *B* factor for the proteins, excluding water molecules and the COOH-terminal tail of α2, is 39.1 Å^2; the average *B* factor of all atoms in the COOH-terminal tail of α2 is 60.4 Å^2.

	Native	IdU^1	IdU^{22}	IdU^{11}
Resolution (Å)	2.74	2.40	2.65	2.40
Measured reflections	74,037	83,570	35,860	76,841
Unique reflections	12,014	16,357	10,715	18,106
Completeness (%)	92.8	90.4	80.0	94.0
Overall $I/\sigma(I)$	9.01	11.7	10.5	14.7
R_{merge} (%)*	13.1	7.4	6.8	5.5
R_{iso} (%)†		11.6	8.2	8.3
R_{cullis}‡		0.57	0.65	0.60
Phasing power§		1.86	0.92	1.82
Mean overall figure of merit (20–2.8Å)‖		0.54		
	Refinement statistics			
*R*esolution range	6–2.8 Å			6–2.5 Å
R factor (%)¶	22.4			22.5
R_{free} factor (%)#	29.2			29.8
Refined geometry	*Overall*	*Protein*	*DNA*	
rmsd bond length (Å)	0.018	0.016	0.019	
rmsd bond angle (°)	1.93	1.83	2.02	

*$R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma I$, *I*: observed intensity, $\langle I\rangle$: average intensity of multiple observations of symmetry-related reflections. †$R_{iso} = \Sigma||F_{pH} - F_p|/\Sigma F_p$, F_{pH} and F_p are the observed derivative and native structure factor amplitude. ‡$R_{cullis} = \Sigma||F_{pH} \pm F_p| - F_{H(calc)}|/\Sigma|F_{pH} - F_p|$. §Phasing power = $\{\Sigma|F_{pH(calc)}|^2/\Sigma[||F_{pH(obs)}| - F_{p(calc)}||]^2\}^{1/2}$, $||F_{pH(obs)}| - F_{p(calc)}|$ is the lack-of-closure error. ‖Mean figure of merit = $\langle|\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)|\rangle$, α: phase, $P(\alpha)$: phase probability distribution. ¶$R = \Sigma||F_p| - |F_{p(calc)}||/\Sigma F_p$. #*R* factor for a subset of 10 percent of the reflection data that were not included in the crystallographic refinement.

NH_2-terminal arm (residues 0 to 7), helices 1 (residues 10 to 23), 2 (residues 28 to 37), and 3 (residues 42 to 58), and the COOH-terminal tail (59 to 74). An additional five COOH-terminal residues, which are disordered in the structure, are not required for a1/α2 repression (*18*). Of the a1 protein, residues 8 to 56 are included in the model. The first 12 amino acids in the protein, including the NH_2-terminal arm, are disordered, as is the COOH-terminal residue of helix 3, Lys^{57}. The a1 and α2 homeodomains, which are 21 percent identical in sequence, are similar in structure and superpose with a 1.0 Å root-mean-square difference in Cα positions (residues 8 to 57).

DNA structure and crystal packing. The DNA in the a1/α2–DNA complex contains a marked overall bend of 60° (Figs. 1A and 3). In contrast is the nearly straight DNA in the complex of α2 alone bound to DNA as discussed below, which contains a bend of 8.8° within a single homeodomain binding site and an overall bend of 7°. Our observations are in agreement with solution studies that have shown that the a1/α2 heterodimer introduces a bend in the DNA estimated at 100°, while α2 alone does not (*21*). The bend in the a1/α2 binding site occurs without dramatic local distortion or kinking of the B-DNA helix. Rather, the DNA helix is smoothly bent, most noticeably at the center of the DNA fragment and in the a1 half of the binding site. The bend is largely the result of a variation in base roll, which adopts negative values near the center of the DNA site and positive values in flanking base pairs (Fig. 3A). A consequence of the bend is a narrowing of the minor groove at the center of the DNA fragment, between the a1 and α2 proteins (Fig. 3A). The continued bending of the DNA in the a1 half of the site leads to a widening of the minor groove and a narrowing of the major groove. The DNA in the ternary complex bends toward the a1/α2 heterodimer, facilitating interactions between the two proteins. The COOH-terminal tail of α2 spans the gap between the a1 and α2 homeodomains at the point where the minor groove of the DNA helix is at its narrowest. Without the observed bend in the DNA, the tail of α2 could not reach its binding site on the back of the a1 homeodomain. The protein-protein interactions at the heterodimer interface, as well as the

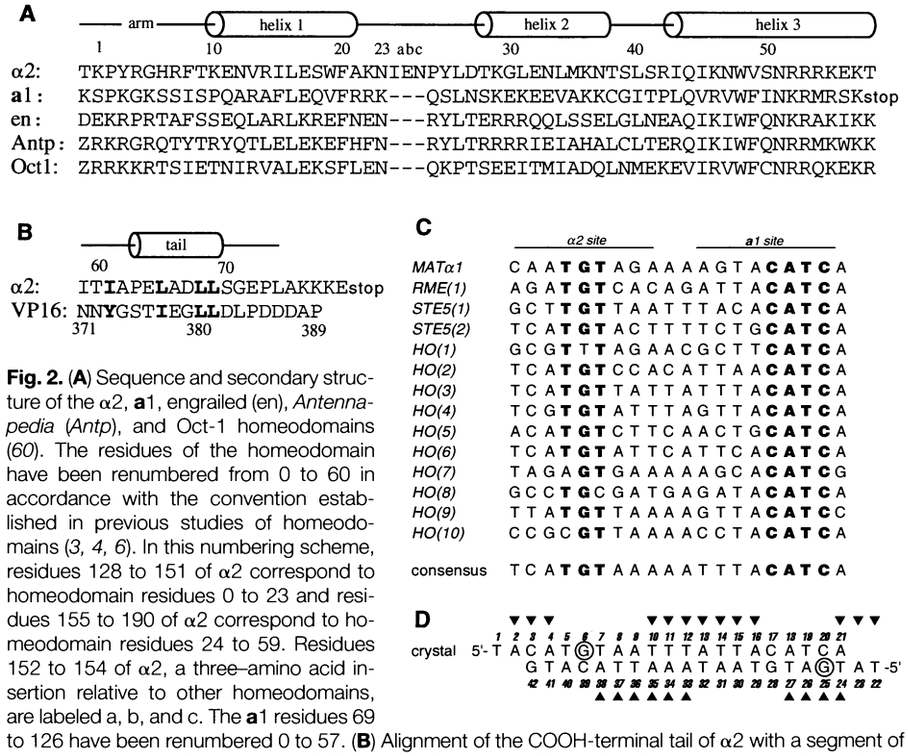
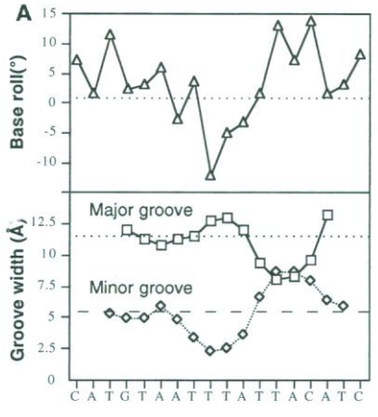


Fig. 2. (**A**) Sequence and secondary structure of the α2, **a**1, engrailed (en), *Antennapedia* (*Antp*), and Oct-1 homeodomains (*60*). The residues of the homeodomain have been renumbered from 0 to 60 in accordance with the convention established in previous studies of homeodomains (*3, 4, 6*). In this numbering scheme, residues 128 to 151 of α2 correspond to homeodomain residues 0 to 23 and residues 155 to 190 of α2 correspond to homeodomain residues 24 to 59. Residues 152 to 154 of α2, a three–amino acid insertion relative to other homeodomains, are labeled a, b, and c. The **a**1 residues 69 to 126 have been renumbered 0 to 57. (**B**) Alignment of the COOH-terminal tail of α2 with a segment of the herpes virus activator protein, VP16, which interacts with the Oct-1 homeodomain. The COOH-terminal tail of α2, residues 191 to 210, have been renumbered 60 to 79. Absolute residue numbers for VP16 are shown below the amino acid sequence. (**C**) Naturally occurring and synthetic DNA binding sites of **a**1/α2. DNA sequences are written 5′ to 3′; invariant and highly conserved bases are highlighted in bold. Sequences are shown of 14 in vivo binding sites for **a**1/α2 located upstream of haploid-specific genes (*16*); the consensus sequence is derived from the list shown. (**D**) Sequence of the double-stranded oligonucleotide used in this crystallographic study and results of in vitro chemical protection experiments (*16*). Circled guanine bases are protected from methylation by binding to DNA of the **a**1/α2 heterodimer. Triangles indicate bases that are protected from hydroxyl radical attack by **a**1/α2.



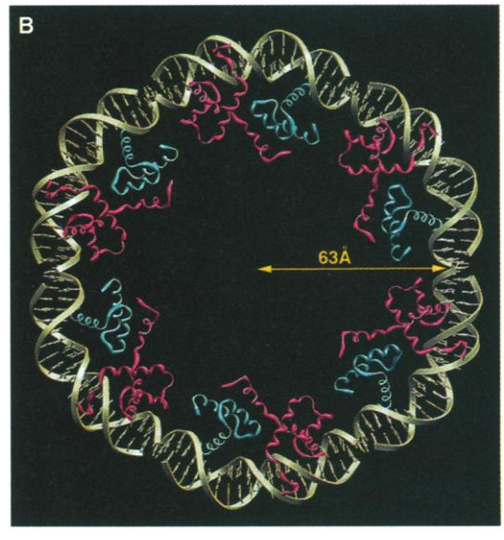


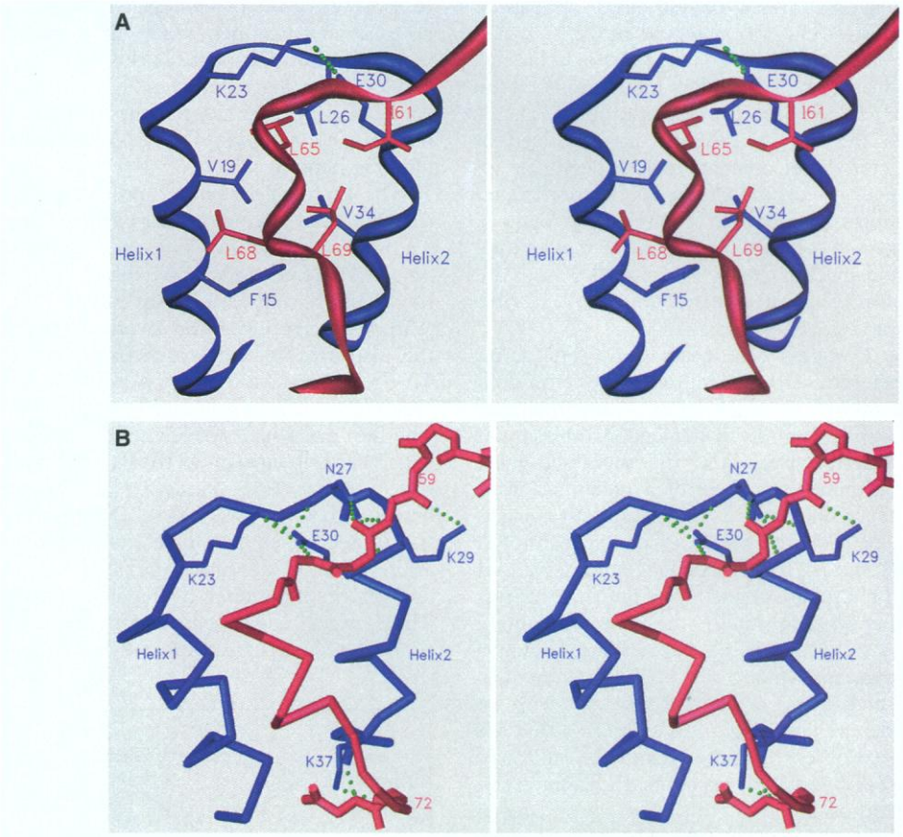
Fig. 3. (**A**) Plot of DNA parameters. A plot is shown of base roll, major groove width, and minor groove width as a function of DNA sequence; horizontal dotted or dashed lines indicate average values for B-DNA. The sequence of the top strand only is shown. Groove widths are measured as the minimum in phosphate-phosphate distance less 5.8 Å for the van der Waals radii of the phosphate groups. All parameters were calculated with the program, CURVES (*61*). (**B**) Packing of complexes in the **a**1/α2–DNA crystals; view looking down the crystallographic 6_1 axis. Adjacent complexes stack end-to-end in the crystal, with the dinucleotide 5′-ApT overhang in one complex forming base-pairing interactions with the complementary 5′-overhanging bases in the DNA of the adjacent complex. With each complex containing a 21-bp DNA fragment that has an overall bend of 60°, the stacked complexes form a superhelical spiral with a radius of 63 Å. The pitch of the superhelix, corresponding to the crystallographic *c* axis, is 45.71 Å.

contacts between **a**1 and α2 and their respective DNA subsites (both described below), probably play a role in stabilizing the bend. In addition, we observed a spine of hydration in the minor groove of the DNA, where it narrows and base roll angles are negative, and in the major groove of the DNA, where bending results in a minimum in major groove width and a maximum positive value of base pair roll angles (Fig. 3A). This hydration may contribute to the stability of the bent DNA conformation, in addition to participating in water-mediated hydrogen bonds between the proteins and the DNA bases.

The DNA bending gives rise to an unusual packing of complexes in the crystal. As has been observed in other crystals of protein-DNA complexes (*4*, *6*, *22*–*26*), the DNA stacks end-to-end in the crystal, forming a pseudocontinuous helix with Watson-Crick base pairing between the overhanging 5′-ApT at the end of one complex and the complementary unpaired bases at the end of the adjacent complex. In the **a**1/α2–DNA crystals, each successive complex bends in the same direction, resulting in a superhelical spiral of complexes that obeys the 6_1 screw symmetry of the space group (Fig. 3B). As measured from the projection shown in Fig. 3B, the radius of curvature of the DNA in the ternary complex is 63 Å. There are crystal contacts between the α2 homeodomain in one complex and the **a**1 protein in the neighboring complex that may limit the degree of bending that can occur at the ends of the DNA. The observed 60° bend may therefore represent an underestimate of the bend induced by **a**1/α2 in a single binding site embedded in a longer fragment of DNA.

Heterodimer interface. The COOH-terminal tail of α2 is unfolded in the monomer, even when bound to DNA. On interaction with the **a**1 protein, the tail of α2 folds to form a complementary surface to its binding site on the **a**1 homeodomain. The heterodimer is stabilized primarily by hydrophobic interactions, in addition to the presence of several hydrogen bonds (Fig. 4, A and B). A hydrophobic patch on the **a**1 homeodomain is formed by Val^{19} of helix 1, Val^{34} of helix 2, and Leu^{26} in the strand that connects helices 1 and 2. These three side chains are partly exposed and lie at the floor of a depression in the surface of the **a**1 homeodomain (Fig. 4, A and C). This depression is flanked at one end by a salt bridge between Lys^{23} and Glu^{30}, and at the other by Phe^{15}.

In the tail of α2, residues 63 to 69 adopt a helical conformation with three leucine side chains (Leu^{65}, Leu^{68}, and Leu^{69}) projecting from one face of the helix (Fig. 4A). In addition, Ile^{61} packs against Leu^{65} and Leu^{69} and helps to stabilize the short helix. The amphipathic helix binds to the **a**1 homeodomain, inserting the three leucine residues into the hydrophobic patch between helices 1 and 2 of **a**1 (Fig. 4, A and C). The **a**1/α2 heterodimer is further stabilized by hydrogen bonds with the **a**1 homeodomain that are mediated by main chain atoms in the COOH-terminal tail of α2 flanking the amphipathic helix (Fig. 4B). The two antiparallel helices of **a**1 and the helix contributed by the tail of α2 form a three-helix bundle, with the helix in α2 parallel to helix 2 of **a**1. The total buried surface area is 754 $Å^2$.



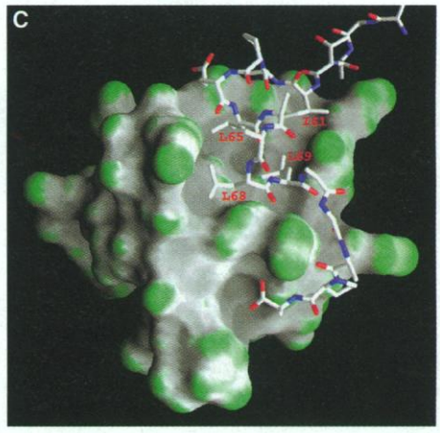


Fig. 4. (**A**) Stereo view of the **a**1/α2 heterodimer interface. The **a**1 protein is shown in blue and the α2 protein is shown in red; helix 3 of the **a**1 homeodomain has been deleted for clarity. In the **a**1 homeodomain, residues Phe^{15}, Val^{19}, Leu^{26}, and Val^{34} form a hydrophobic patch between helices 1 and 2. Residues 63 to 69 in the tail of α2 form a distorted amphipathic helix with residues Leu^{65}, Leu^{68}, and Leu^{69} exposed on one face. Ile^{61}, located NH_2-terminal to the amphipathic helix, packs against Leu^{65} and Leu^{69}. The hydrophobic face of the α2 tail forms favorable van der Waals contact with the hydrophobic patch between helices 1 and 2 of the **a**1 homeodomain, as well as with the salt bridge formed by Lys^{23} and Glu^{30}. (Additional contacts, not shown, are formed with the aliphatic side chain of Arg^{22}.) (**B**) Stereo view of the hydrogen-bonding interactions between **a**1 and α2. The main chain NH of Ala^{62} in α2 donates a hydrogen bond to the Oε of Glu^{30} in **a**1, and the second Oε accepts a hydrogen bond from the peptide NH of Asn^{27} of **a**1. The side chain of Asn^{27} in **a**1 forms a bridging contact, donating a hydrogen bond to the peptide N of Thr^{60} in α2 and capping helix 2 of **a**1 by accepting a hydrogen bond from the peptide N of Lys^{29}. The side chain of Lys^{29} donates a hydrogen bond to the carbonyl of Ile^{59}. COOH-terminal to the amphipathic helix in the α2 tail, the carbonyl of Glu^{72} hydrogen bonds with the Nζ of Lys^{37} in **a**1, while the amide N of Glu^{72} forms a hydrogen bond with the carbonyl O of Lys^{37}. (**C**) Depiction of the molecular surface of the **a**1 homeodomain, with a stick model of the COOH-terminal tail of α2. The **a**1 surface is color-coded such that the most convex part of the surface is green, the most concave part is gray, and planar surfaces are white [the figure was made with the program GRASP (*62*)].

The results of mutagenesis studies of the α2 protein verify the importance of the heterodimer contacts we observed. Point mutations that change a wild-type residue to alanine were introduced at various positions in the α2 tail and assayed for their effect on the ability of the intact α2 and **a**1 proteins to repress transcription in yeast (Fig. 5). The strongest effect of alanine substitution was observed at Ile^{61}, Leu^{65}, Leu^{68}, and Leu^{69}, the four hydrophobic residues that mediate key interactions in the **a**1/α2 complex. Substitution of side chains in the α2 tail that are not involved in heterodimer contacts have a negligible effect on repression (Fig. 5). In prior studies substitution of Leu^{65} with Ser was found to disrupt the ability of **a**1/α2 to repress haploid-specific genes in vivo (*27*) and the affinity of intact **a**1/α2 for DNA in vitro was 200 times lower (*28*).

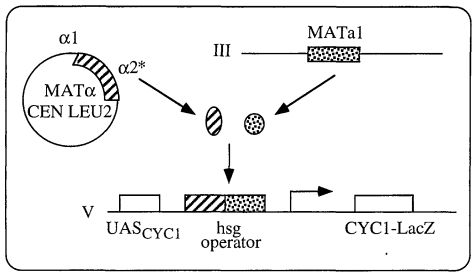
Protein-DNA interactions. The **a**1/α2 heterodimer forms an extensive set of contacts with a DNA binding site that spans 18 bp (Fig. 6A). Each homeodomain contacts both the bases and the sugar-phosphate backbone by means of a combination of direct side chain contacts and water-mediated hydrogen bonds. Common features of **a**1 and α2 binding include a conserved set of phosphate contacts and one base contact that serve to stabilize the homeodomain on its binding site. A key residue is Asn^{51}, invariant among all homeodomains (*2*), which forms a bidentate contact with an adenine base (Fig. 6, A, B, and D) that has been observed in the engrailed, Oct-1, and α2-DNA complexes (*4*, 6, 8). Flanking this base, **a**1 and α2 contact four phosphates in the same way (Fig. 6A). These contacts are mediated by four residues that are identical in both proteins—Leu^{26}, Gln^{44}, Trp^{48}, and Arg^{53}—and by the main chain NH of residue 8. DNA sequence recognition presumably arises from contacts with the DNA bases that are mediated by side chains that differ between the two homeodomains.

The α2 protein contacts a total of 7 bp in either the major or minor groove. Four base pairs in the major groove are contacted by three side chains in helix 3: Asn^{51}, Arg^{54}, and Ser^{50} (Fig. 6B). Of these, Ser^{50}, which contributes to DNA binding specificity differences among homeodomains (*29–31*), forms water-mediated hydrogen bonds with 2 bp, and Arg^{54} donates hydrogen bonds to a base (Gua^{6}) and to a side chain, Asn^{51}, which contacts Ade^{38}. The Arg^{54} contact accounts for the observation that Gua^{6} is the sole guanine in the α2 site that is protected from methylation by **a**1/α2 binding. In the minor groove, 5 bp are contacted by three residues in the NH_2-terminal arm of α2—Arg^{4}, Gly^{5}, and Arg^{7} (Fig. 6C). Two of the base pairs contacted in the minor groove are also contacted by side chains in the major groove, thereby accounting for the high DNA sequence conservation at these positions (Fig. 2C). In addition to contacting 2 bp, Arg^{7} hydrogen bonds to a water molecule that is part of the spine of hydration observed where the minor groove narrows. Further stabilization of the α2 homeodomain on its binding site is provided by a set of hydrogen bonds, salt bridges, and van der Waals interactions with the sugar-phosphate backbone of the DNA that are diagrammed in Fig. 6A. Most of these arise from side chains in helix 3 that form contacts with the DNA backbone flanking helix 3.

The **a**1 homeodomain is positioned on the DNA in a manner similar to that of the α2 protein, contacting bases in the major groove with five residues in helix 3—Val^{47}, Ile^{50}, Asn^{51}, Met^{54}, and Arg^{55} (Fig. 6D). With the exception of the invariant Asn^{51}, the identities of DNA-contacting residues 47, 50, 54, and 55 differ from those in α2 and hence mediate different base contacts. The Ile^{50} and Met^{54} residues are in van der Waals contact with, respectively, Thy^{15} and Cyt^{17}. Also, Arg^{55} forms two hydrogen bonds with Gua^{26}, thereby accounting for the protection of this base from methylation when **a**1/α2 binds to DNA; Arg^{55} donates an additional hydrogen bond to Thy^{19}, which is base-paired with the adenine (Ade^{26}) that is contacted by the invariant Asn^{51}. The Val^{47} residue forms additional van der Waals contacts with Ade^{26}. Furthermore, 10 residues participate in contacts with the DNA backbone (Fig. 6A). The observed major groove contacts are consistent with the sequence conservation in the **a**1 subsite (Fig. 2C): All four invariant DNA positions are contacted in the major groove and two of the base pairs (Thy^{19}·Ade^{26} and Cyt^{20}·Gua^{25}) are contacted by more than one side chain.

In addition to direct side chain–DNA contacts, we see a network of five water molecules immobilized in the major groove at the interface between helix 3 and the DNA (Fig. 6D). These water molecules participate in hydrogen bond interactions with a total of 4 bp, and are stabilized by a hydrogen bond formed by one of the water molecules with the guanidinium of Arg^{46}, which also forms a salt bridge with a phosphate. Formation of this network of water-mediated contacts requires the observed local DNA curvature. An extensively hydrated interface between helix 3 and the DNA has also been observed by solution NMR in the Antennapedia-DNA complex (*32*).

Unlike that of the α2 homeodomain, the NH_2-terminal arm of **a**1 is disordered in the crystal. This result was unanticipated, as our present structure of the α2 homeodomain and earlier structural studies of homeodomains (*4–6*, *8*) show that the NH_2-



	β-gal units	Pheromone Production a-factor	Pheromone Production α-factor	Mating Behavior as **a**	Mating Behavior as α
vector only	45.0	+++	--	+	--
MAT α2	5.4	+	--	+	--
T58A	6.5	+	--	+	--
I59A	7.0	+	--	+	--
T60A	6.3	+	--	+	--
I61A	22.7	+	+++	+	+
P63A	7.4	+	--	+	--
E64A	7.8	+	--	+	--
L65A	44.1	+	+++	+	+
D67A	6.1	+	--	+	--
L68A	19.5	+	+++	+	+
L69A	32.4	+	+++	+	+
E72A	5.9	+	--	+	--

Fig. 5. Effects of point mutations in the COOH-terminal tail of α2 on **a**1/α2–mediated repression in vivo. Alanine was substituted for the wild-type residue at various positions in the COOH-terminal tail of α2. Low copy (CEN) plasmids containing the MATα locus with specific site-directed mutations in the MATα2 gene (α2* in the figure) were transformed into a MAT**a** strain and repression by **a**1/α2 was monitored: (i) by expression of a test promoter whose expression is controlled by a single hsg operator, (ii) by the ability of transformants to produce mating pheromones, and (iii) by the ability of the transformants to mate. The mutations in the COOH-terminal tail of α2 are shown in the first column on the left. The second column shows the expression of a CYC-LacZ reporter construct that contains an hsg operator. In a cell carrying a wild-type MATα2 gene on a plasmid, the β-galactosidase reporter gene is expressed at approximately one-ninth the level of a cell containing the plasmid vector only [this level of repression is not complete, probably because of plasmid loss—see (*28*)]. Four of the α2 tail mutants fail to repress efficiently the test promoter, with Leu^{65}→Ala having the strongest effect. The third and fourth columns show the production of mating pheromones by the transformants. For all of the mutants, production of **a** factor is repressed as compared to the vector control, indicating that the α2 mutant proteins are synthesized in the cell and can still function with MCM1 to repress the **a**-specific genes (*18*). The fourth column indicates that the four mutants that show significant derepression of the test promoter fail to repress STE12, a haploid-specific gene required for α-factor production. The fifth and sixth columns show the mating behavior of the transformants. The phenotype of the four mutants that enable the yeast to mate as α cells are consistent with a defect in **a**1/α2–mediated repression.

terminal arm of the homeodomain binds in the minor groove of the DNA. Chemical modification experiments that probe the protection of the DNA backbone from hydroxyl radical attack in the presence of the **a**1/α2 heterodimer are consistent with the binding of the NH_2-terminal arm of **a**1 in the minor groove (*16*). It is possible that the NH_2-terminal arm of **a**1 has been displaced from the minor groove in the **a**1/α2–DNA crystal, where the NH_2-terminal arm of **a**1 would be expected to bind near the junction between two double-stranded synthetic oligonucleotides. The 5′ phosphate group that would be present at that junction in a continuous DNA strand could be of importance in stabilizing binding of the NH_2-terminal arm of **a**1 to DNA. If there are indeed further DNA contacts formed with bases in the minor groove, some may contribute further to the observed DNA sequence preference of the **a**1 protein.

Because of the extensive set of contacts formed by the **a**1 homeodomain with DNA, it is surprising that **a**1 does not bind DNA detectably on its own. There are several possible explanations for this phenomenon. One is that DNA bending is required for the protein-DNA contacts we observe. The energy to induce and stabilize this bend might be too great for **a**1 alone, but may be facilitated by the presence of α2 on the DNA and the interaction between the α2 tail and the **a**1 homeodomain. Another explanation may be that the binding of the tail of α2 to **a**1 induces a small conformational change in the **a**1 homeodomain that increases its overall DNA binding energy. Solution NMR studies of the **a**1 homeodomain have shown that, on heterodimerization with α2, the resonances in **a**1 that are most perturbed on complex formation cor-

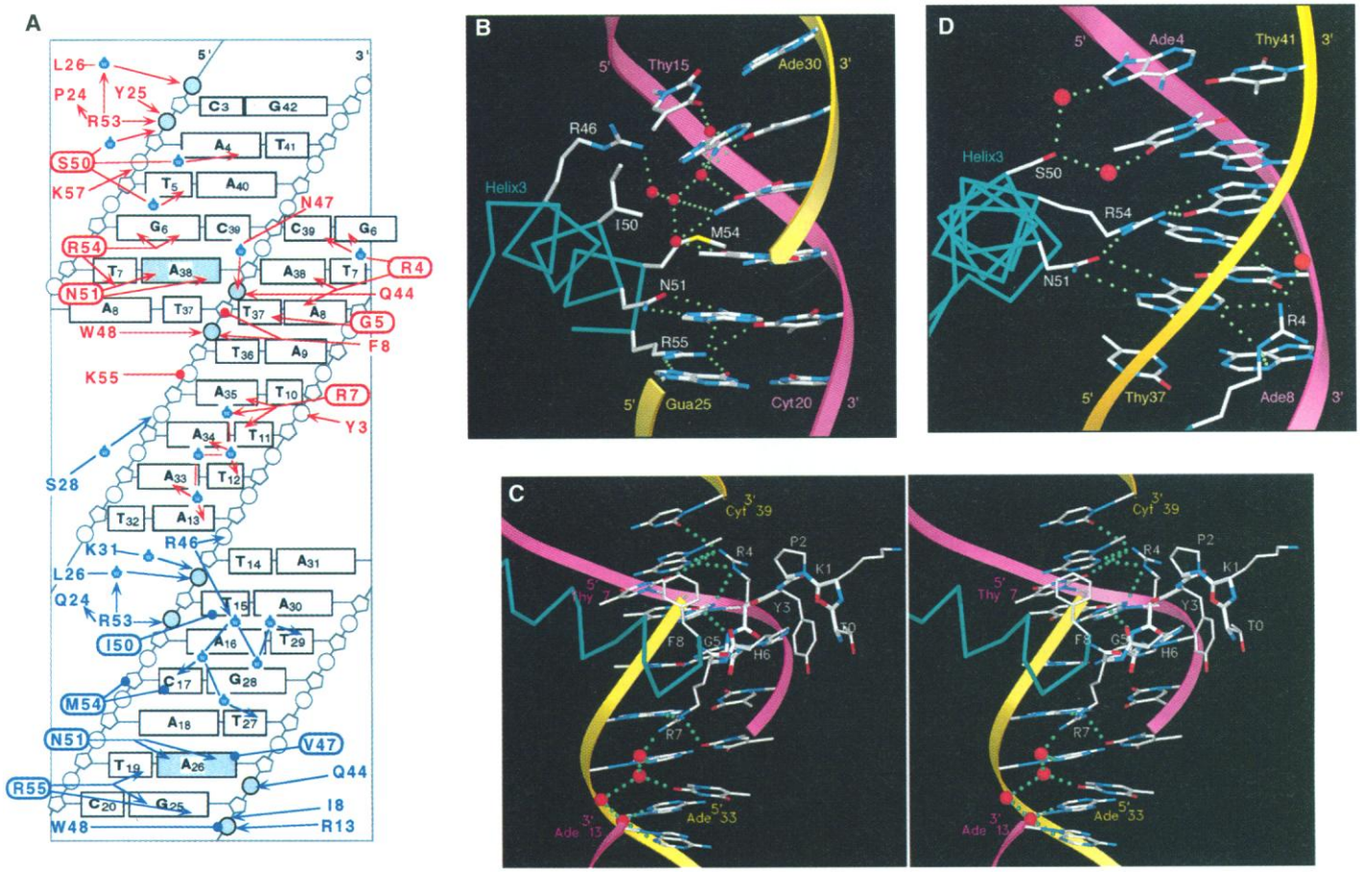


Fig. 6. Protein-DNA interactions. (**A**) Schematic diagram summarizing contacts formed by both **a**1 and α2 with the DNA. The α2-mediated contacts are shown in red; **a**1-mediated contacts are shown in blue. Residues encircled with ovals form base contacts. Hydrogen bonds and salt bridges are indicated by arrows; van der Waals interactions are indicated by lines that end in circles. Bond criteria are as follows: for hydrogen bonds, less than or equal to 3.3 Å separation of a donor-acceptor pair; van der Waals interactions, less than 4.0 Å distance between contacting groups. Residues involved in base contacts are encircled with ovals. The shaded phosphate groups and adenine bases are contacted in a conserved manner by both **a**1 and α2. The complex equilibrium constant for dissociation of the ternary complex into free **a**1, α2, and DNA is 10^{-15} M^2 for the **a**1 and α2 fragments used in this study and 10^{-16} M^2 for the intact proteins (*11, 20*). (**B**) Contacts formed between helix 3 of α2 and bases in the major groove. Asn^{51} contacts Ade^{38}, with the Oδ of the side chain accepting a hydrogen bond from the N6 of the base and whose Nδ is in contact with the N7 (dist. = 3.4 Å). Arg^{54} hydrogen bonds to the N7 and O6 of Gua^6 with its NH1, while its NH2 group forms a bridging hydrogen bond to the Oδ of Asn^{51}. Ser^{50} forms water-mediated hydrogen bonds with the N7 of Ade^4 and the O4 of Thy^5. (**C**) Stereo view of contacts between the NH_2-terminal arm of α2 and bases in the minor groove of the DNA. A total of five bases are contacted by three side chains in the NH_2-terminal arm of α2. Arg^4 contacts 2 bp directly and a third via water-mediated hydrogen bonds. The Nε of Arg^4 is within hydrogen bonding distance of the N3 of Ade^8 and the N3 of Ade^{38}, whereas the NH_2 contacts the Gua^6·Cyt^{39} base pair via water-mediated hydrogen bonds. The peptide NH of Gly^5 donates a hydrogen bond to the O2 of Thy^{37}, while the guanidinium group of Arg^7 donates hydrogen bonds to the N3 of Ade^{35} and the O2 of Thy^{11}. In addition, the NH_2 of Arg^7 helps to stabilize the spine of hydration in the minor groove. (**D**) Contacts between the **a**1 homeodomain and bases in the major groove. Ile^{50} is in van der Waals contact with the methyl of Thy^{15}, while Asn^{51} forms two hydrogen bonds with Ade^{26}, whose Oδ accepts a hydrogen bond from the N6 of the base and whose Nδ donates a hydrogen bond to the N7. The Cγ of Met^{54} is in van der Waals contact with the C6 of the Cyt^{17} pyrimidine, while the Cε is in van der Waals contact with the adjoining subdeoxyribose. The Arg^{55} side chain forms contacts with two base pairs, its NH1 donating a hydrogen bond to the N7 of Gua^{25} and its NH2 donating hydrogen bonds to the O6 of Gua^{25} and the O4 of Thy^{19}. In addition, there is a network of five bound water molecules in the major groove that hydrogen bond to the guanidinium of Arg^{46}, as well as 3 bp. Not shown in this depiction is Val^{47}, whose Cγ is in van der Waals contact with the C8 of Ade^{26}.

respond to residues that lie in portions of helices 1 and 2 and in the loop connecting them (*33*). A likely set of contacts that could be affected by a conformational change in the loop are those mediated by the invariant homeodomain residue Arg^{53}, which participates in an extensive set of contacts with the sugar-phosphate backbone, a bound water molecule, and the backbone of **a**1 in the loop between helices 1 and 2 (Fig. 6A) (*34*). Very small changes in the loop conformation could easily disrupt this network of contacts, thereby diminishing the affinity of **a**1 for DNA in the absence of α2. Since Arg^{53} is invariant among all homeodomains and mediates a conserved set of contacts in **a**1, α2, and other homeodomains, it is likely to be crucial for formation of favorable interactions between **a**1 and its DNA binding site.

In previous studies, the homeodomain was found to be positioned on the DNA in a highly conserved manner by DNA contacts that are mediated by conserved residues (*4*–*6*, *8*). A pertinent question is whether the orientation of the homeodomain on the DNA and the contacts it forms are perturbed as a result of heterodimerization. We find that, despite the additional protein-protein interactions and the concomitant distortion in the DNA, the α2 homeodomain in the heterodimer binds DNA in a manner essentially identical to that found in the structure of α2 alone bound to DNA (Fig. 7). This similarity extends to the side chain contacts formed by α2 with its binding site and the local structure of the DNA. The few differences between the two structures are most likely the result of differences in the resolution of the structure determinations (*35*). Because of the similarity in the way α2 is positioned on the DNA in the presence and absence of **a**1, it is very surprising that a mutant α2 protein with Ala substituted for Ser^{50}, Asn^{51}, and Arg^{53} exhibits no significant difference in the affinity of the **a**1/α2 heterodimer for DNA while greatly reducing the affinity of α2 alone for DNA (*28*). It is possible, however, that this triple mutation reduces the DNA sequence discrimination of the **a**1/α2 heterodimer without impairing its affinity for the DNA.

The structure of the **a**1/α2–DNA complex shows how the binding specificity of a homeodomain can be raised by complex formation with a second protein. The heterodimerization of the α2 and **a**1 homeodomains, each with only modest affinity and specificity for DNA, results in a heterodimer that binds DNA in a highly specific manner, preferring its own binding site over random DNA by a ratio of at least 10^5 (*11*). The specificity of this interaction derives from at least three sources: the specificity of the interactions between the COOH-terminal tail of α2 and the **a**1 homeodomain, summation of the DNA sequence preferences of the individual proteins, and the precise binding site spacing imposed by the nature of the heterodimer interface. It is clear from the **a**1/α2–DNA structure that insertion or deletion of base pairs between the **a**1 and α2 binding sites would disrupt heterodimer contacts, and this observation is supported by studies of **a**1/α2 binding to altered DNA sites (*16*, *35a*). A final contributing factor to ternary complex stability may be the relative deformability of a given sequence of DNA in that DNA bending is required for formation of the complex that we observe. Thus the sequence at even noncontacted bases may contribute to overall complex stability. The net result of all these considerations is a larger set of criteria that must be met for a stable **a**1/α2–DNA complex to form, and hence a much higher specificity of binding than is observed for the monomeric α2 or **a**1 proteins.

Multi-protein complex formation in transcription. The work presented above shows how the interaction between two homeodomain proteins can be mediated by a flexible tail that becomes ordered on complex formation. There are now numerous examples of cooperative interactions involving homeodomain proteins, some of which might be mediated by the same types of interactions observed in the **a**1/α2 heterodimer. The *Caenorhabditis elegans* homeodomain protein MEC-3 heterodimerizes with the UNC-86 POU domain protein, which also contains a homeodomain (*36*). This interaction is reminiscent of **a**1/α2 heterodimer formation in that it is dependent upon the 16 amino acids COOH-terminal to the MEC-3 homeodomain. Cooperative interactions have been observed between the *Drosophila* extradenticle (exd) homeodomain protein and the Ultrabithorax (Ubx), engrailed, and abdominal-A homeodomain proteins (*37*–*38*). These interactions require either NH_2- or COOH-terminal extensions to the homeodomain, one as short as 15 residues (*37*). Pbx1, a human homeodomain protein closely related to exd (*39*), similarly has been observed to bind DNA cooperatively with several Hox homeodomain proteins (*40*). The Pbx-Hox interactions are dependent on a short NH_2-terminal peptide in the Hox proteins and a COOH-terminal tail in Pbx. The homeodomain most closely related to both Pbx and exd is **a**1, with which they share 40 percent sequence identity (*41*); in contrast, **a**1 and α2 share only 21 percent sequence identity. The sequence similarity between **a**1, Pbx, and exd may imply the existence of a common mechanism used by these proteins to recognize their homeodomain partners.

The asymmetric nature of **a**1/α2 heterodimer formation raises the possibility that a similar type of interaction could occur between a homeodomain protein and a nonhomeodomain protein containing a peptide similar to the tail of α2. An example of this may be the mammalian Oct-1 POU homeodomain protein, which binds DNA cooperatively with the herpes virus VP16 transcriptional activator (*42*). Point mutations in Oct-1 that disrupt the Oct-1–VP16 complex are located in helices 1 and 2 of the Oct-1 homeodomain. As shown above and noted in the NMR study of **a**1/α2 (*33*), the location of the Oct-1 mutations is analogous to the region on the **a**1 homeodomain that is contacted by the tail of α2 (*43*, *44*). If Oct-1 is acting as the analog of **a**1, what part of VP16 may play a role similar to the COOH-terminal tail of α2? The results of deletion and mutagenesis studies have identified a 12-residue region of VP16 responsible for interaction with the Oct-1 homeodomain (*45*, *46*). This region (residues 376 to 387), which has been proposed to form an amphipathic helix, contains sequence similarity to the helix in the α2 tail that contacts **a**1 (Fig. 2B). Model-

Fig. 7. Interaction of α2 with DNA in the **a**1/α2–DNA ternary complex as compared with the structure of α2 alone bound to DNA. The two structures were aligned by performing a least-squares superposition of the α2 homeodomain in the **a**1/α2–DNA complex with one of the α2 homeodomains in the structure of α2 alone bound to DNA (*6*). Since the latter structure contains two α2 homeodomain monomers bound to a 21-bp DNA fragment, the homeodomain chosen for the alignment is the one bound to the identical 9-bp sequence to which α2 is bound in the ternary complex. The second α2 homeodomain contained in the α2-DNA structure, which does not contact the other α2 monomer, has been omitted for clarity. The conformation of the α2 homeodomains, the local structure of the DNA, and the protein-DNA contacts are nearly identical in the two complexes. There are pronounced differences outside the α2 binding sites due to the overall 60° bend in the DNA induced by the **a**1/α2 heterodimer.

building studies show that this peptide from VP16, when folded in the conformation of the α2 tail, could pack between helices 1 and 2 of Oct-1 in a manner similar to that observed in the a1/α2 heterodimer.

It is likely that there are transcription factors from other structural classes that are also bound in protein-protein complexes by a flexible peptide that adopts a distinctive conformation only upon complex formation. There are strong parallels between the manner in which α2 interacts with a1 and the way in which it forms a complex with the MADS box protein (*47*), MCM1 (*48*, *49*). A short region NH_2-terminal to the α2 homeodomain that is unstructured in the free protein is required for cooperative binding with MCM1 (*50*). This NH_2-terminal peptide specifies complex formation with MCM1, as splicing it to the engrailed homeodomain confers on engrailed the ability to bind DNA cooperatively with MCM1 (*50*). Thus α2 has evolved as a DNA binding protein capable of interaction with two structurally distinct partners by acquiring peptide extensions both NH_2- and COOH-terminal to its DNA binding domain that specify complex formation with different partners. The presence of flexible protein-recognition peptides that extend from stably folded DNA binding domains may prove to be a general feature of the architecture of other classes of eukaryotic transcriptional regulators.

REFERENCES AND NOTES

1. W. J. Gehring *et al.*, *Cell* **78**, 211 (1994).
2. A. Laughon, *Biochemistry* **30**, 11357 (1991).
3. Y. Q. Qian *et al.*, *Cell* **59**, 573 (1989).
4. C. R. Kissinger, B. Liu, E. Martin-Blanco, T. B. Kornberg, C. O. Pabo, *ibid.* **63**, 579 (1990).
5. G. Otting *et al.*, *EMBO J.* **9**, 3085 (1990).
6. C. Wolberger, A. K. Vershon, B. Liu, A. D. Johnson, C. O. Pabo, *Cell* **67**, 517 (1991).
7. C. L. Phillips, A. K. Vershon, A. D. Johnson, F. W. Dahlquist, *Genes Dev.* **5**, 764 (1991).
8. J. D. Klemm, M. A. Rould, R. Aurora, W. Herr, C. O. Pabo, *Cell* **77**, 21 (1994).
9. M. Affolter, A. Percival-Smith, M. Müller, W. Leupin, W. J. Gehring, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4093 (1990).
10. T. Hoey and M. Levine, *Nature* **332**, 858 (1988).
11. C. Goutte and A. D. Johnson, *J. Mol. Biol.* **233**, 359 (1993).
12. A. D. Johnson, in *Transcriptional Regulation*, S. L. McKnight, K. R. Yamamoto, Eds. (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1992), vol. 2, pp. 975.
13. J. C. W. Shepherd, W. McGinnis, A. E. Carrasco, E. M. De Robertis, W. J. Gehring, *Nature* **310**, 70 (1984).
14. C. Goutte and A. D. Johnson, *Cell* **52**, 875 (1988).
15. A. M. Dranginis, *Nature* **347**, 682 (1990).
16. C. Goutte and A. D. Johnson, *EMBO J.* **13**, 1434 (1994).
17. M. R. Stark and A. D. Johnson, *Nature* **371**, 429 (1994).
18. A. Mak and A. D. Johnson, *Genes Dev.* **7**, 1862 (1993).
19. C.-Y. Ho, J. G. Adamson, R. S. Hodges, M. Smith, *EMBO J.* **13**, 1403 (1994).
20. C. Phillips, M. R. Stark, A. D. Johnson, F. W. Dahlquist, *Biochemistry* **33**, 9294 (1994).
21. D. L. Smith, A. B. Desai, A. D. Johnson, *Nucleic Acids Res.* **23**, 1239 (1995).
22. A. K. Aggarwal, D. W. Rodgers, M. Drottar, M. Ptashne, S. C. Harrison, *Science* **242**, 899 (1988).
23. J. A. Anderson, M. Ptashne, S. C. Harrison, *Nature* **326**, 846 (1987).
24. S. C. Schultz, G. C. Shields, T. A. Steitz, *Science* **253**, 1001 (1991).
25. C. Wolberger, Y. Dong, M. Ptashne, S. C. Harrison, *Nature* **335**, 789 (1988).
26. J.-A. Feng, R. C. Johnson, R. E. Dickerson, *Science* **263**, 348 (1994).
27. J. Strathern, B. Shafer, J. Hicks, C. McGill, *Genetics* **120**, 75 (1988).
28. A. K. Vershon, Y. Jin, A. D. Johnson, *Genes Dev.* **9**, 182 (1995).
29. S. D. Hanes and R. Brent, *Science* **251**, 426 (1991).
30. A. Percival-Smith, M. Mueller, M. Affolter, W. J. Gehring, *EMBO J.* **9**, 3967 (1990).
31. J. Treisman, P. Gonczy, M. Vashinshtha, E. Harris, C. Desplan, *Cell* **59**, 553 (1989).
32. M. Billeter *et al.*, *J. Mol. Biol.* **234**, 1084 (1993).
33. S. M. Baxter, D. M. Gontrum, C. L. Phillips, A. F. Roth, F. W. Dahlquist, *Biochemistry* **33**, 15309 (1994).
34. Arg^{53} donates two hydrogen bonds to phosphate P4 via its Nε and NH2 groups. The NH1 group donates one hydrogen bond to the main chain carbonyl of residue 24 and a second hydrogen bond to a bound water molecule. The water, which is found in an identical position in both the a1 and α2 subsites, in turn accepts a hydrogen bond from the peptide NH of Leu^{26} and donates a hydrogen bond to phosphate P3.
35. As compared with the structure of α2 alone bound to DNA, the α2 homeodomain in the heterodimer contains five additional ordered residues in the NH_2-terminal arm, some of which mediate minor groove contacts. In addition, no water molecules were placed in the α2-DNA structure, which was determined at a resolution of 2.7 Å.

35a. Y. Jin, J. Mead, T. Li, C. Wolberger, A. K. Vershon, *Science* **270**, 290 (1995).

36. D. Xue, Y. Tu, M. Chalfie, *ibid.* **261**, 1324 (1993).
37. F. B. Johnson, E. Parker, M. A. Krasnow, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 739 (1995).
38. S.-K. Chan, L. Jaffe, M. Capovilla, J. Botas, R. S. Mann, *Cell* **78**, 603 (1994); M. A. van Dijk and C. Murre, *ibid.*, p. 617.
39. J. Nourse *et al.*, *ibid.* **60**, 535 (1990); M. P. Kamps, C. Murre, X. Sun, D. Baltimore, *ibid.*, p. 547; C. Rauskolb, M. Peifer, E. Wieschaus, *ibid.* **74**, 1101 (1993).
40. C.-P. Chang *et al.*, *Genes Dev.* **9**, 663 (1995).
41. W. A. Flegel *et al.*, *Mechan. Dev.* **41**, 155 (1993).
42. S. Stern, M. Tanaka, W. Herr, *Nature* **341**, 624 (1989).
43. J. L. Pomerantz, T. M. Kristie, P. A. Sharp, *Genes Dev.* **6**, 2047 (1992).
44. J. S. Lai, M. A. Cleary, W. Herr, *ibid.*, p. 2058.
45. R. F. Greaves and P. O'Hare, *J. Virol.* **64**, 2716 (1990).
46. S. Stern and W. Herr, *Genes Dev.* **5**, 2555 (1991).
47. H. Ma, M. F. Yanofsky, E. M. Meyerowitz, *ibid.*, p. 484.
48. C. A. Keleher, C. Goutte, A. D. Johnson, *Cell* **53**, 927 (1988).
49. R. T. Sauer, D. L. Smith, A. D. Johnson, *Genes Dev.* **2**, 807 (1988).
50. A. K. Vershon and A. D. Johnson, *Cell* **72**, 105 (1993).
51. T. Li, M. Stark, A. D. Johnson, C. Wolberger, *Proteins* **21**, 161 (1995).
52. Crystals were flash-frozen and maintained at −179°C, x-ray diffraction data were collected with an *R*-AXIS II image plate detector mounted on a Rigaku RU-200 rotating anode x-ray generator, with CuKα radiation filtered by a graphite monochromator. The reflection data were processed, merged, and scaled with the *R*-AXIS software (Molecular Structure). Subsequent phase refinement and map calculations were performed with the PHASES suite of programs [W. Furey and S. Swaminathan, *American Crystallographic Association Meeting Abstracts* PA33 (1990)]. Iodine positions were located by Patterson and difference Fourier methods, and heavy atom parameters were calculated and refined to produce a set of MIR phases.
53. B. C. Wang, *Methods Enzymol.* **115**, 90 (1985).
54. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta. Crystallogr.* **A47**, 110 (1991).
55. A. T. Brünger, *Nature* **355**, 472 (1992).
56. _____, *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR* (Yale Univ Press, New Haven, 1992).
57. R. J. Read, *Acta. Crystallogr.* **A46**, 900 (1990).
58. A. Hodel, S.-H. Kim, A. T. Brünger, *ibid.* **A48**, 851 (1992).
59. S. V. Evans, *J. Mol. Graph.* **11**, 134 (1993).
60. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
61. R. Lavery and H. Sklenar, *J. Biomol. Struct. Dyn.* **6**, 63 (1988).
62. A. Nicholls, K. Sharp, B. Honig, *Proteins*, **11**, 281 (1991).

63. We thank M. Glover, J. Kim, M. Rould, M. Bianchet, A. Gittis, J. Klemm, and W. Weis for advice and suggestions and M. Amzel, D. Leahy, and J. Berg for comments on the manuscript. Time on a Cray supercomputer was provided by the National Cancer Institute–Frederick Biomedical Supercomputing Center. Supported by NSF grant MCB-9304526 (C.W.), the David and Lucile Packard Foundation (C.W.), and the Lucille P. Markey Charitable Trust (C.W.) and by a National Institutes of Health grant GM-37049 (A.D.J.). The coordinates have been deposited in the Protein Data Bank in Brookhaven, NY, with accession number 1YRN.

26 July 1995; accepted 13 September 1995