Cooperative Ordering in Homeodomain−DNA Recognition: Solution Structure and Dynamics of the MAT**a**1 Homeodomain†

Janet S. Anderson,‡ Michael D. Forman,§,|| Sarah Modleski,§ Frederick W. Dahlquist,⊥ and Susan M. Baxter*,§

Department of Chemistry, Union College, Schenectady, New York 12308, Wadsworth Center and Department of Biomedical Sciences, New York State Department of Health, Empire State Plaza, P.O. Box 509, Albany, New York 12201, and Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403

Received March 24, 2000; Revised Manuscript Received June 13, 2000

ABSTRACT: The mating type homeodomain proteins, MAT**a**1 and MATα2, combine to form a heterodimer to bind DNA in diploid yeast cells. The **a**1−α2 heterodimer tightly and specifically binds haploid-specific gene operators to repress transcription. On its own, however, the **a**1 homeodomain does not bind DNA in a sequence-specific manner. To help understand this interaction, we describe the solution structure and backbone dynamics of the free **a**1 homeodomain. Free **a**1 in solution is an ensemble of structures having flexible hinges at the two turns in the small protein fold. Conformational changes in the **a**1 homeodomain upon ternary complex formation are located in the loop between helix 1 and helix 2, where the C-terminal tail of α2 binds to form the heterodimer, and at the C-terminus of helix 3, the DNA recognition helix. The observed differences, comparing the free and bound **a**1 structures, suggest a mechanism linking van der Waals stacking changes to the ordering of a final turn in the DNA-binding helix of **a**1. The tail of α2 induces changes in loop 1 of **a**1 that push it toward a properly folded DNA binding conformation.

Transcription factors containing the homeodomain DNA-binding motif have been found in eukaryotic species from yeast to human (*1*). The homeodomain consists of 60 amino acids that fold into three helices with an unstructured N-terminal arm. Crystallographic (*2*−*8*) and NMR[1] spectroscopic (*9*−*11*) studies of homeodomain−DNA complexes have revealed a general model of DNA recognition (*12*) among homeodomains involving DNA contacts in the major and minor grooves via the third helix and the N-terminal arm, respectively. Homeodomain proteins bind with relatively low sequence specificity and affinity in vitro (*13*). Despite their conserved protein fold, their mode of DNA binding, and their modest affinity for DNA, homeodomain proteins achieve a range of extremely specific regulatory actions that lead to precise developmental outcomes in vivo.

To achieve the specific functionality that leads to tightly regulated gene expression in vivo, homeodomain motifs are often flanked by adjacent amino acid sequences that mediate protein−protein interactions, providing combinatorial advantages for recognition and affinity. For instance, the HOX proteins contain a characteristic YPWM sequence found N-terminal to their homeodomain which has been shown (*7*, *8*) to mediate protein−protein interactions in both the human Pbx1−HoxB1 and the *Drosophila* Ubx−Exd complexes. The *Saccharomyces cerevisiae* homeodomain MATα2, or α2, has two regions flanking its homeodomain for protein−protein interactions important in entirely separate functions. In both α and **a**−α yeast cell types, an N-terminal sequence of α2 interacts with the MCM1 protein, leading to repression of **a**-specific haploid genes (*14*−*16*). In diploid **a**−α cells, the C-terminal tail of α2 binds the MAT**a**1, or **a**1, homeodomain to form a heterodimer that recognizes haploid-specific gene (*hsg*) operators (*4*, *17*−*19*). The **a**1−α2 heterodimer recruits the SSN6−TUP1 complex, which represses transcription of the haploid-specific genes (*20*). These genetic, biochemical, and structural studies show that homeodomains can use specific protein−protein combinations to achieve functional versatility.

The yeast **a**1 protein, however, is distinct among well-studied homeodomains. Whereas the α2 monomer alone can bind DNA with modest affinity, the **a**1 monomer shows no sequence-specific DNA binding in vitro (*18*, *19*). But in combination, the **a**1−α2 heterodimer is 3000 times more specific for *hsg* operators than for nonspecific DNA (*18*). Recent studies (*21*, *22*) have shown that **a**1 provides the majority of the DNA specificity and affinity for the **a**1−α2 heterodimer. Indeed, Vershon and co-workers (*23*) have shown that an α2 variant, with alanine substituting for residues that make base-specific DNA contacts, can still repress transcription in complex with **a**1.

The **a**1 protein has a homeodomain motif at its C-terminus with no adjacent, flanking sequences for protein−protein interactions identified to date. The C-terminal segment of the protein (**a**1_{66-126}) has been identified as the minimal fragment that is able to heterodimerize and bind tightly to

† This work was supported by a grant from the National Institute of General Medical Sciences (GM55361) to S.M.B.

* To whom correspondence should be addressed. Phone: (518) 486-7320. Fax: (518) 473-2900. E-mail: baxter@wadsworth.org.

‡ Union College.

§ New York State Department of Health.

|| Present address: Pfizer Central Research, Box 111W, Eastern Point Road, Groton, CT 06340.

⊥ University of Oregon.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single-quantum coherence; en, engrailed homeodomain; Antp, Antennapedia homeodomain.

10.1021/bi000677z CCC: $19.00
Published on Web 07/28/2000

DNA, with only a 10-fold reduction in DNA affinity relative to that of the intact protein (*19*). Previous NMR (*24*) and crystallographic studies (*4*) demonstrate that this minimal, functional fragment of the **a**1 protein has a classical homeodomain fold. The ternary complex (**a**1−α2−*hsg* operator) structure (*4*) demonstrates that the **a**1 homeodomain makes extensive DNA contacts using its third helix, but the N-terminal arm is disordered in the crystal and does not make the minor groove contacts characteristic of homeodomain proteins. Structural studies have shown that the **a**1−α2 heterodimer is formed both in the absence (*24*) and in the presence of DNA (*4*). Heterodimerization involves the formation of a helix in the C-terminal tail of α2 that contacts a hydrophobic patch on the surface of the **a**1 homeodomain. Recently, Stark et al. (*22*) have demonstrated that even the isolated α2 C-terminal tail peptide can convert **a**1 from a weak to a strong DNA-binding protein. This finding suggests that the changes in the **a**1 homeodomain dictated by the α2 C-terminal tail underlie the high affinity and specificity of the **a**1−α2 heterodimer for its DNA targets.

Using NMR spectroscopy, we have determined the solution structure and described the backbone dynamics of the functional **a**1 homeodomain fragment, $\mathbf{a}1_{66-126}$. Free **a**1 in solution may be most accurately described as an ensemble of structures having flexible hinges at the two turns in the small protein fold. When free **a**1 is compared with the bound form determined crystallographically, conformational changes upon ternary complex formation are located in the loop between helix 1 (H1) and helix 2 (H2), where the C-terminal helical peptide of α2 binds, and at the C-terminus of the third helix (H3), the DNA recognition helix. The interaction of the α2 tail causes rearrangements of side chains in the loop structure of **a**1. The effects of these changes may reduce the flexibility in the **a**1 homeodomain structure, effectively reducing the entropic barrier for subsequent DNA binding by the **a**1−α2 heterodimer.

MATERIALS AND METHODS

NMR Measurements. The **a**1 homeodomain was overexpressed and purified as previously described (*25*). NMR data were acquired at 25 °C on Bruker Avance DRX 500 and 600 MHz spectrometers. Uniformly labeled ([U-^{15}N] or [U-^{15}N,^{13}C]) protein samples were dissolved in 25 mM deuterated acetate (pH 4.5), 100 mM KCl, and 0.01% NaN_3. DSS was added as the internal chemical shift standard. Spectra were processed using FELIX 970 (Molecular Simulations, Inc.). Complete backbone and 97% of the side chain proton resonance assignments for the **a**1 homeodomain are supplied in the Supporting Information and will be deposited in the BioMagResBank at University of Wisconsin.

For the assignment of NOEs, ^{1}H−^{15}N NOESY-HSQC, methyl−methyl ^{13}C−^{13}C NOESY (*26*), and two-dimensional NOESY spectra, collected at 100 ms, were used. WATERGATE pulses (*27*) were used for water suppression. The three-dimensional NOESY-HSQC experiment was recorded as a 1024 × 256 × 96 complex matrix, with eight scans and a recycle time of 1.5 s. The methyl−methyl NOESY experiment was recorded as a 1024 × 152 × 38 complex matrix, with 16 scans and a recycle time of 1.5 s. The two-dimensional experiment was collected as a 4096 × 1024 complex matrix, with 32 scans and a recycle time of 1.5 s.

A series of ^{1}H−^{15}N correlation spectra were collected at 500 MHz to determine the backbone ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rate constants and the steady-state NOE. Pulse sequences used to measure relaxation rates have been described previously (*28*) using the PEP-Z modification. For R_1 experiments, relaxation delays of 17.3 (×2), 207 (×2), 345, 518, 777, and 2500 (×2) ms were used (duplicate acquisitions noted by ×2). For R_2 experiments, relaxation delays of 8.14 (×2), 16.29 (×2), 24.43 (×2), 32.58 (×2), 40.71, 57.01, 73.30, 146.59 (×2), 171.02, 219.88, and 472.35 (×2) ms were used. The interval between the refocusing pulses in the ^{15}N CPMG sequence was 1 ms. ^{1}H−^{15}N steady-state NOE values were determined from interleaved two-dimensional spectra recorded with and without a 3 s presaturation pulse in addition to a 3 s relaxation delay, resulting in a 6 s recycle time. Spectra were recorded in a 2048 × 256 complex matrix. For each t_1 increment, 32 scans were acquired. For all spectra, an exponential line broadening function was applied in ω_2; shifted sine bell functions were used to apodize data in ω_1.

Relaxation Data Analysis. Individual peak heights were measured using FELIX macros written by M. Akke (Lund University, Lund, Sweden) and supplied by A. G. Palmer (Columbia University, New York, NY). Uncertainties in peak heights were estimated from duplicate spectra as described previously (*28*). R_1 and R_2 values were extracted using nonlinear, least-squares fitting of the experimental peak heights versus time using monoexponential functions provided in software packages by A. G. Palmer (Columbia University) or in Kaleidagraph (Synergy Software).

The model-free formalism was used to relate the relaxation parameters to the spectral density function of the NH bond vector. The relaxation rates, R_1 and R_2, for an amide ^{15}N nucleus are dominated by dipolar interaction with the directly attached proton and by chemical shift anisotropy (*29*):

$$R_1 = (d^2/4)[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N) \quad (1)$$

$$R_2 = (d^2/8)[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + (c^2/6)[4J(0) + 3J(\omega_N)] + R_{ex} \quad (2)$$

for which the constants $d = (\mu_0 h \gamma_H \gamma_N / 8\pi^2)\langle r_{NH}^3 \rangle$ and $c = \omega_N \Delta\sigma/\sqrt{3}$, where μ_0 is the permittivity of free space, h is Plank's constant, γ_H and γ_N are the gyromagnetic ratios of the proton and the nitrogen nucleus, respectively, ω_H and ω_N are the ^{1}H and ^{15}N Larmor frequencies, respectively, r_{NH} is the internuclear ^{1}H−^{15}N distance (1.02 Å), and $\Delta\sigma$ (−160 ppm) is the chemical shift anisotropy measured for ^{15}N nuclei by Hiyama et al. (*30*) Exchange contributions to R_2 due to conformational averaging on the microsecond to millisecond time scale are treated by adding a third term, R_{ex}, to eq 2.

Analysis of the relaxation data within the model-free formalism was performed using the Modelfree program (version 3.1), kindly provided by A. G. Palmer, and described in previously published procedures (*31*). Initial estimates of the overall correlation time, τ_m, were obtained from the 10% trimmed mean of the R_1/R_2 ratio. Model selection was carried out as described by Mandel et al. (*31*).

Structure Calculations. To compare our results with the current homeodomain literature, the amino acid residues

of **a**1 were renumbered from the previous system of 66−126 (*24*) to −3 to 57. A preliminary set of NOEs was assigned from both ^{15}N-edited and homonuclear NOESY spectra. Five categories were used: strong (1.8−2.7 Å), medium (1.8−3.5 Å), weak (1.8−5.0 Å), very weak (1.8−6.0 Å), and extremely weak (1.8−8.0 Å for methyl−methyl NOEs or NOEs involving aromatic HD* or HE* protons). For $^3J_{HNH\alpha}$ values of >8, torsion angle restraints for the dihedral angle ϕ were set equal to $-120 \pm 60°$. For $^3J_{HNH\alpha}$ values of <6, ϕ was calculated using the Karplus relation with constants *A*, *B*, and *C* being equal to 6.51, −1.76, and 1.60, respectively, and the restraint was given an uncertainty of ±30°. No restraints were used for residues with intermediate $^3J_{HNH\alpha}$ values. Hydrogen bond restraints were included for those identified by slow hydrogen exchange rates in earlier work (*24*).

This set of 615 NOEs, 38 dihedral angle restraints, and 42 hydrogen bond restraints were used in an X-PLOR (*32*) refinement to produce a cluster of medium-resolution structures. A representative structure of this cluster was selected as a starting conformation for an abbreviated ARIA (*32*) calculation where any remaining unassigned peaks from the two-dimensional and three-dimensional data sets and proton chemical shift assignments described above were calibrated and assigned in a single iteration of ARIA. Both the unambiguous and the ambiguous ARIA assignments were examined for possible inclusion in subsequent X-PLOR calculations using the following criteria. Any ARIA assignment contributing at least 40% to the peak intensity and having a distance between assigned protons of ≤7 Å was included in an additional group of NOEs to be added to the previous list of NOEs.

This enlarged set of NOEs and the dihedral angle and hydrogen bond restraints were cycled through several X-PLOR calculations. After each calculation, NOE violations were identified with the program AQUA (*33*). Any NOE which was violated by an average of ≥0.10 Å was reduced in strength or deleted from the set if no further reductions in strength were possible. The final set of 1178 NOE connectivities consisted of 283 intraresidue, 350 sequential, 380 medium-range, and 165 long-range NOEs, or an average of 24.0 NOE restraints per structured residue (residues 8−56). Structure refinement was based on unambiguously assigned NOEs, 42 hydrogen bonds, and 38 scalar coupling constants ($^3J_{HNH\alpha}$).

One hundred random structures were generated in X-PLOR using the parallhdg.pro parameter file supplied with ARIA as it had more precise values of bond lengths and bond angles than the original parameter file supplied with X-PLOR, and standard distance geometry and simulated annealing calculations were performed. The 100 final structures were analyzed with the program NMRCLUST (*34*). Sixty-six conformations formed a cluster with a rmsd of 0.338 Å for backbone atoms of residues 11−53. The 20 lowest-energy conformations of the 100 final structures were selected to represent the final ensemble of solution structures of **a**1. These structures had an rmsd of 0.264 Å and target function values ranging from 114.4 to 129.2 kcal/mol. NMRCLUST also identified one of the 20 to be the most representative of the cluster of low-energy conformations. The structure coordinates (1F43) and NMR constraint files are deposited at the Protein Data Bank.

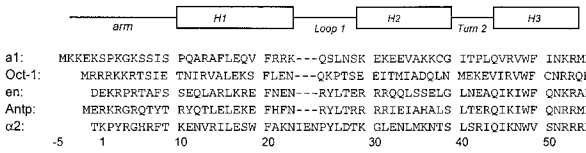

FIGURE 1: Sequence and approximate secondary structure alignment of the **a**1, Oct-1, engrailed (en), Antennapedia (Antp), and MATα2 homeodomains. The C-terminal **a**1 homeodomain fragment (residues 66−126) has been renumbered −3 to 57 in accordance with previous studies of homeodomains and the **a**1−α2 system (*42*).

Table 1: Number of Constraints Used in Structure Determination of MAT **a**1 Homeodomain and Structural Statistics

type of restraint	no. of restraints
total NOE restraints	1178
intraresidue	283
sequential	350
medium-range ($\|i-j\| \leq 4$)	380
long-range ($\|i-j\| > 4$)	165
dihedral angle restraints	39
hydrogen bond restraints	42
NOEs per structured residue (residues 8−56)	24
NOE violations of > 0.25 Å	none
average NOE violation	0.03 Å
backbone rmsd for the 20 lowest-energy structures (residues 11−53)	0.264 Å
all heavy atom rmsd for 20 lowest-energy structures (residues 11−53)	0.963 Å

RESULTS

Resonance Assignments. The **a**1 homeodomain was expressed and purified using the *dnaY* system described previously (*25*). The amino acid sequence is shown in Figure 1 and is aligned with the sequences of the Oct-1, engrailed (en), Antennapedia (Antp), and MATα2 homeodomains. Mass spectroscopic analysis confirmed that the [U-99.5%^{15}N] and [U-95%^{15}N,^{13}C] proteins, including an N-terminal methionine, were correctly expressed and free of post-translational modifications or oxidation. The majority of backbone amide chemical shift assignments was reported previously (*24*) for the folded residues of the **a**1 homeodomain. Complete sequence-specific assignments for all amide protons in the protein construct, residues −3 to 57 (Figure 1), coexpressed with *dnaY* (*25*) are reported in the Supporting Information. In addition, 97% of the side chain protons for residues 9−57 (fragment numbering scheme) have been assigned. For many cases, methylene protons for the longer chain amino acids are overlapped.

Structure Calculations. From the NMR data sets, 615 NOEs, 38 dihedral angle restraints, and 42 hydrogen bond restraints were generated for initial model building. To expand the set of input data, we used the ARIA program (*35*) to help assign ambiguous NOE resonances as described in Materials and Methods. The final set (Table 1) of 1178 NOEs consisted of 283 intraresidue, 350 sequential, 380 medium-range, and 165 long-range NOEs, or an average of 24 NOE restraints per structured residue (residues 8−56). This enlarged set of NOE connectivities, along with the dihedral angle and hydrogen bond restraints reported previously (*24*), was cycled through several X-PLOR (*32*) calculations to refine the three-dimensional structure. The sequence folds into three helices separated by a loop and a turn. The N-terminal arm residues (Lys 3−Ser 7) are unstructured as are residues 54−57 at the C-terminal end of

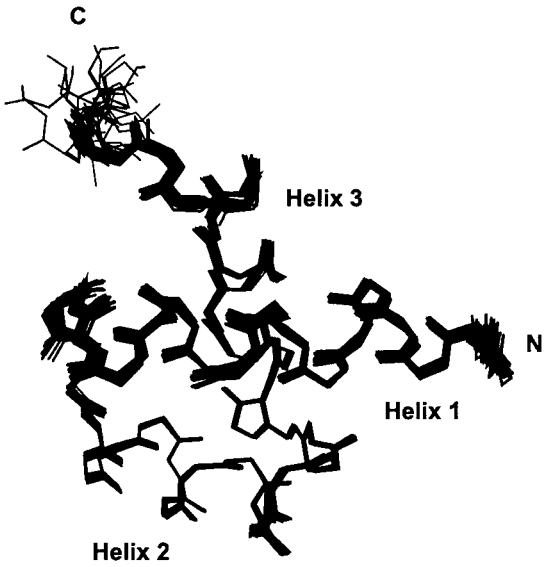

FIGURE 2: Superposition of 20 selected and refined solution structures of the **a**1 homeodomain fragment (residues 8−57). Unstructured residues Lys −3 to Ser 7 are omitted for clarity. Labels indicate each helix and the N- and C-termini of the protein.

helix 3. Figure 2 shows the superposition of the 20 lowest-energy conformations of **a**1. The solution structure calculations result in a family of 20 structures with an average backbone rmsd of 0.264 Å.

The quality of the set of 20 final structures was assessed with the AQUA and PROCHECK NMR programs (*33*). The Ramachandran plot showed that 99.9% of structured residues 8−53 had ϕ and ψ values in the allowed regions. The structure selected by NMRCLUST to be representative of the 20 lowest-energy conformations was analyzed using the secondary structure definition method of Kabsch and Sander (*36*). Thus, H1 is defined as extending from Pro 10 to Arg 22, H2 from Ser 28 to Cys 38, and H3 from Gln 44 to Lys 52. Using this Kabsch and Sander analysis, both the loop and the turn of the homeodomain consist of six residues. No consensus structure was determined for the N-terminal residues of the homeodomain, from Lys −3 to Ser 9.

The most probable loop conformers are captured in the solution structure refinement. An analysis of root-mean-square standard deviations of side chain heavy atoms (not shown) and circular variance in the 20 representative solution structures of the **a**1 homeodomain highlights the variation in positions of side chains in the H1 and loop 1 regions compared to the helix−turn−helix segment of the homeodomain. Indeed, residues in this region have a below average completeness of NOEs as judged by the AQUA program (*37*). Specifically, the circular variance (CV), or the spread of dihedral angles, for the Phe 20 side chain ($CV_{\chi^2} = 0.371$) is significantly higher than for Phe 24 ($CV_{\chi^2} = 0.001$) and Phe 53 ($CV_{\chi^2} = 0.007$) that are better packed in the protein core.

To verify the solution structure of **a**1, 10% of the NOEs used in the final structure determination were removed at random. Fifty random starting conformations were generated and the X-PLOR calculations performed as described in Materials and Methods. The helix angles of the 10 new lowest-energy structures were unchanged from their previous values, although the uncertainty increased. This procedure was repeated after removing a total of 20% of the NOEs, with the same result.

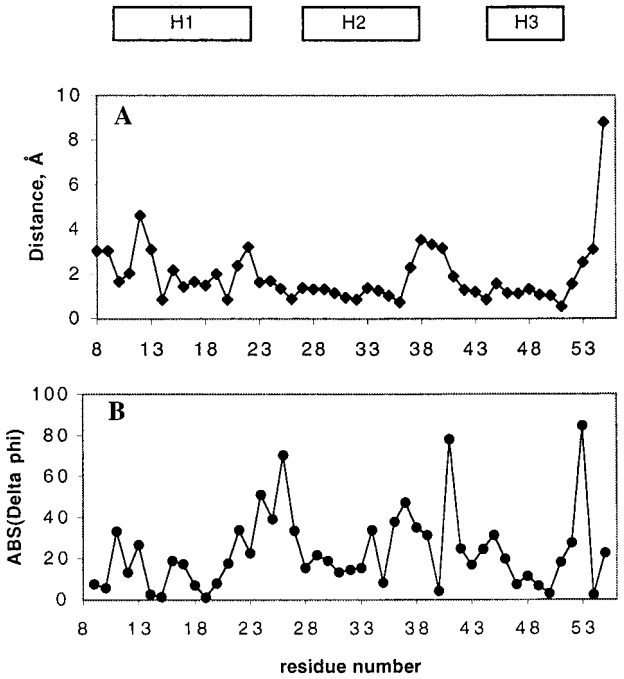

FIGURE 3: (A) C^{α} backbone displacements comparing **a**1 free in solution and bound in a ternary complex (**a**1−α2−*hsg* operator) plotted vs residue number. The two structures were matched from residues 10−52 using the MidasPlus program (*54*, *55*) giving a rmsd of 1.83 Å over the backbone N, C^{α}, and C′ atoms. (B) Plot of the magnitude of differences in ϕ backbone torsion angle between **a**1 free in solution and bound in the ternary complex plotted vs residue number.

*Comparison of the Solution Structure of **a**1 with the Bound Form in the **a**1−α2−DNA Complex.* With the solution structure of the **a**1 homeodomain in hand, we compared the free structure to the bound form of the protein as found in the crystal structure of the **a**1−α2−*hsg* operator complex (*4*). Although the side chain atoms of Gln 18, Arg 21, Glu 32, and Ser 56 were not resolved in the crystal structure, backbone comparisons should be valid. Figure 3A shows the distance between corresponding α-carbons for the representative NMR structure and the X-ray structure of **a**1 in the ternary complex (*4*) resulting in a rmsd of 1.83 Å over the backbone atoms. The largest deviation appears at the C-terminus, where the third helix is unraveled from residues 53 to 55 compared with the helix in the ternary complex. Other differences in backbone displacement are significantly smaller but tend to occur in the loop and turn regions of the homeodomain.

Examination of the backbone dihedral angles can also indicate differences between two conformations of the same protein, with the advantage that analysis of dihedral angles is independent of how the structures are matched. Figure 3B shows the magnitude of the differences in the backbone torsion angle ϕ for the solution and X-ray structures. Local differences in backbone conformation are clustered in the loop between H1 and H2, although Thr 41 and Arg 53 also undergo significant ϕ angle changes (Figure 3B). Similar results were observed for the backbone dihedral angle ψ (not pictured).

To further analyze the effects of backbone dihedral changes, the relative orientation of the helices was calculated. An axis for each helix was determined using the program Helix (*38*), and then angles between the helical axes were

Table 2: Comparison of Homeodomain Helix Angles

data collection	structure	H1/H2 angle (deg)	H1/H3 angle (deg)	H2/H3 angle (deg)
NMR structures, **a**1	average, 20 lowest-energy conformations	139 ± 1	84 ± 3	97 ± 2
	average, 10% NOEs removed	138 ± 3	78 ± 6	100 ± 6
	average, 20% NOEs removed	139 ± 3	82 ± 4	94 ± 3
NMR structures, homeodomains	Antp (C39S), free[a]	152	58	110
	Antp (C39S), bound to DNA[b]	163	58	110
	VND/NK-2, free[c]	157	49	110
	VND/NK-2, bound to DNA[d]	155	62	108
	Oct-1, free[e]	149	55	123
	average, 11 structures[f]	156 ± 7	59 ± 5	113 ± 5
X-ray structures, homeodomains	MAT **a**1 in **a**1−α2−*hsg* operator complex[g]	152	60	106
	eng, free[h]	156	59	114
	eng, bound to DNA[i]	160	61	112
	Oct-1, bound to DNA[j]	150	64	112
	average, 21 structures[k]	156 ± 5	60 ± 3	109 ± 5

[a] PDB file 2HOA used for comparison. [b] PDB file 1AHD used for comparison. [c] PDB file 1VND used for comparison. [d] PDB file 1NK3 used for comparison. [e] PDB file 1POG used for comparison. [f] Average of 11 structures from 2HOA, 1AHD, 1VND, 1NK3, 1FTT, 1FTZ, 1HDP, 2LFB, 1OCP, 1POG, and 1BW5 used for comparison. [g] PDB file 1YRN used for comparison. [h] PDB file 1ENH used for comparison. [i] PDB file 3HDD used for comparison. [j] PDB file 1OCT used for comparison. [k] Average of 21 structures from 1YRN, 1ENH, 3HDD, 2HDD, 1AKH, 1APL, 1B72, 1FJL, 1MNM, 1B8I, and 1OCT used for comparison.

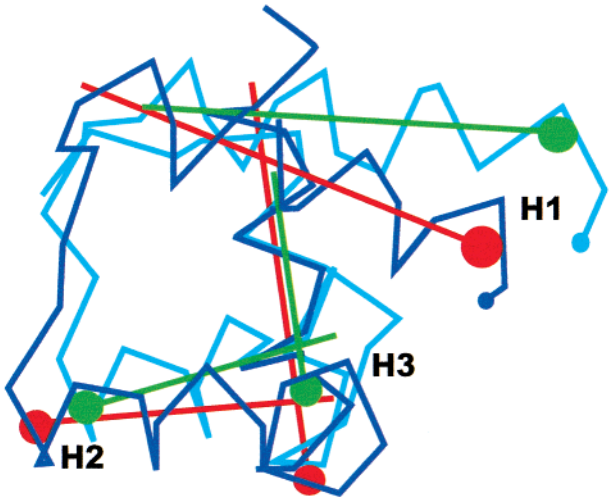

FIGURE 4: Superposition of **a**1 homeodomain structures (residues 9−50) both free in solution (light blue) and bound (dark blue) in the ternary complex (**a**1−α2−*hsg* operator). Only residues Ser 9−Ile 50 were compared to avoid matching the ends of the homeodomain where the largest differences exist. Bars indicate helix axes as calculated by the program Helix (*38*) and plotted with a *Mathmatica* notebook. The solution structure calculations suggest that H1 is oriented differently in the free form of the homeodomain and in the bound form.

computed. Table 2 shows the average of these angles for the 20 lowest-energy solution structures of **a**1 and, as a comparison, for **a**1 in the ternary complex. The angle between H2 and H3 is somewhat different for the free and bound forms of **a**1, but the angles between H1 and H2, and H1 and H3, differ by an amount significantly larger than the uncertainty in the solution values. It is likely that H1 has a different orientation with respect to the rest of the homeodomain (HTH) in the solution structure of **a**1. This is shown directly in Figure 4, where the representative solution structure and the X-ray structure are superimposed.

^{15}N Relaxation Rate Measurements. NMR spectra of the **a**1 homeodomain provide qualitative evidence of a range of backbone dynamics. ^{1}H−^{15}N HSQC spectra contain eight sharp, intense resonances assigned to residues in the N-terminal segment of the sequence. In addition, the resonances for Leu 16, Lys 23, Leu 26, and Thr 41 are doubled with an approximate ratio of major to minor peaks of 8:1. Several other peaks are broadened and weak in intensity relative to other HSQC resonances, including Arg 21, Arg 22, Ser 25, Asn 27, and Cys 38. To further quantify **a**1 homeodomain backbone dynamics, we measured ^{15}N relaxation rates for 51 out of 61 residues in the **a**1 homeodomain. Missing data for assigned residues are due to resonance overlap, especially for residues in the N-terminal segment of the sequence.

The experimental values for relaxation rates, R_1 and R_2, plotted against residue number are shown in Figure 5. The average value of T_1 (=1/R_1) was 461 ± 74 ms; the average T_2 (=1/R_2) was 140 ± 70 ms. The large standard deviations reflect the range of relaxation rates manifested by backbone ^{15}N nuclei in the **a**1 homeodomain. Residues in the N-terminal segment of the homeodomain, characterized by reduced relaxation rates, are unstructured by other criteria, such as the lack of medium- and long-range NOEs, negative ^{15}N NOEs (data not shown), and intermediate values of $^{3}J_{HNH\alpha}$. Similar dynamic behavior is noted for the C-terminal residues, Ser 56 and Lys 57. When relaxation times for residues with negative and very low ^{15}N{^{1}H} NOE values are excluded from calculation of the mean, the average value of T_1 was 433 ± 32 ms; the average T_2 was 120 ± 43 ms. The structured region of the **a**1 homeodomain, encompassing residues 8−56, contains two regions characterized by increased R_2 rates and decreased R_1 rates, relative to the mean. Both segments encompass turns in the **a**1 homeodomain structure, loop 1 residues Lys 23−Asn 27 and turn 2 residues Cys 38−Pro 42. The relaxation data suggest that the rmsd of the backbone structure reported here may suggest more structure and rigidity than the raw NMR data indicate.

Model-Free Analysis. The experimentally determined relaxation rates and heteronuclear ^{15}N{^{1}H} NOE data for structured residues Gln 11−Arg 55 were fit to spectral density functions to calculate order parameters (S^2). All the experimental data were adequately fit to spectral density functions containing one or two adjustable parameters. Once spectral density models were selected, the overall correlation time (τ_m) was simultaneously optimized for all residues. The optimized correlation time for the **a**1 homeodomain was 5.09 ± 0.04 ns.

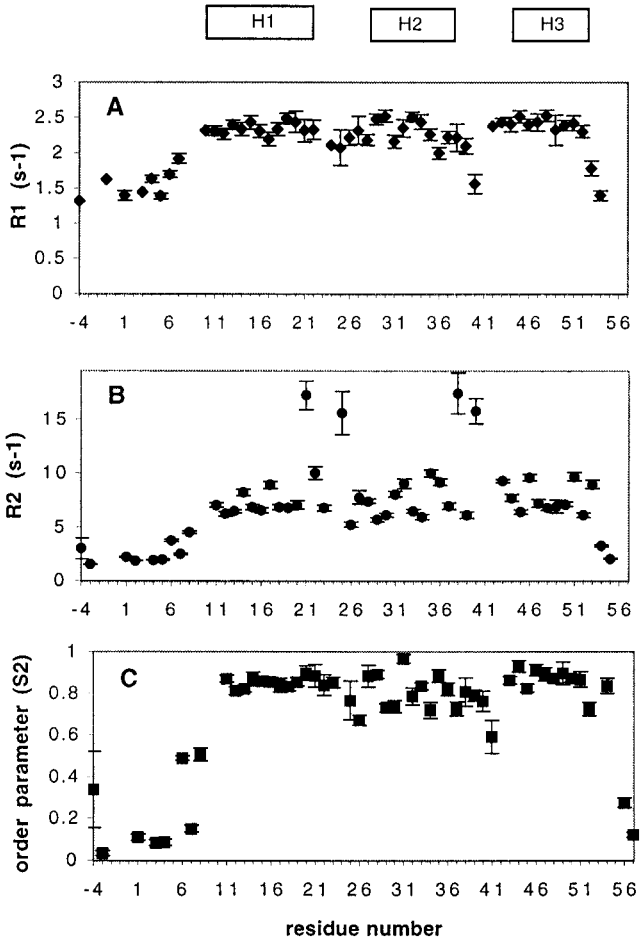

FIGURE 5: Experimental values for ^{15}N relaxation rates, R_1 and R_2 (in s^{-1}), and the order parameters (S^2) determined for the **a**1 homeodomain plotted vs residue number. Measured ^{15}N R_1 rates are shown in panel A and ^{15}N R_2 rates in panel B. Order parameters (S^2) are plotted in panel C. Gaps in the data result from overlapping or broadened resonances that made measurements unreliable. Error bars are shown.

As qualitatively predicted from the relaxation data and line shape analysis, the N-terminal arm residues (from Lys −3 to Ser 9) have very low ($S^2 < 0.5$) order parameters, as do the two C-terminal residues (Figure 5C). The structured homeodomain residues, in both helices and turns, have order parameters averaging 0.83 ± 0.07, a typical value for backbone residues of folded, globular proteins. Only two backbone amides have significantly lower order parameters: Leu 26 in loop 1 (0.67) and Thr 41 in turn 2 (0.59).

The relaxation data for 21 residues of the **a**1 homeodomain were adequately fit by the simple spectral density function. Residues in the N-terminal arm and at the end of H3 required an additional term (effective correlation time, τ_e) to account for fast motions. In addition, some residues in the structured part of the protein required an additional τ_e term, including Leu 26, Lys 29, Glu 30, Val 34, Val 45, and Lys 52. Leu 26 is contained in loop 1. Lys 29, Glu 30, and Val 34 are contained in the first turn of H2; Val 45 is at the beginning of H3. Lys 52 is the last residue in H3.

Fifteen residues required an additional R_{ex} term that suggests local backbone motions on the microsecond to millisecond time scale. Their locations in the **a**1 homeodomain structure are generally at the interfaces between secondary structure elements (the transition from helix to turn, for instance). We also find that residues at the end of the third helix require exchange terms, indicative of helix fraying. Three other residues requiring R_{ex} terms (Ala 14, Glu 17, and Arg 46) are contained in the middle of helices 1 and 3, at the point that the two helices cross, suggesting some flexibility in the interface between these helices. Four residues (Arg 21, Ser 25, Cys 38, and Ile 40) require R_{ex} terms greater than 9 s^{-1}. Arg 21 is the likely C-terminal capping residue for H1, although the structure of the C-terminal helix cap does not fit any of the canonical capping motifs described by Aurora and Rose (*39*). Ser 25, Cys 38, and Ile 40 are all loop and turn residues. All residues requiring R_{ex} terms manifest broadened or doubled resonances in the NMR spectra.

DISCUSSION

The **a**1 homeodomain has been set apart from other homeodomains characterized in the literature because it does not bind DNA on its own in vitro. Nonetheless, previous work has shown that the **a**1 homeodomain contributes a surprising amount of binding energy to the functional repressor complex it forms with the α2 protein (*21*−*23*). The current model of cooperative binding, presented by Stark and co-workers (*22*), is that α2 effects a change in **a**1 that converts **a**1 from a weak to strong DNA-binding protein. Here we report the solution structure of the **a**1 homeodomain, supplemented with backbone dynamics information. Not surprisingly, on the basis of earlier NMR characterization (*24*), the **a**1 structure (Figure 2) consists of three helices connected by two turns, similar to all other homeodomain structures determined to date.

*The N-Terminal Arm of **a**1 Behaves Differently from That of Other Homeodomains.* The N-terminal residues (from Lys −3 to Gly 4) of the **a**1 homeodomain are largely unstructured ($S^2 < 0.2$) and do not contact the folded, helical domain in any significant or specific manner in solution. Unlike the Oct-1 (*40*) and Isl-1 (*41*) homeodomains, **a**1 exhibits no unambiguous NOE connectivity between the N-terminal arm and the folded helical core. Chemical shift mapping of **a**1 titrated with α2 (*24*) indicates that heterodimerization does not affect or involve the N-terminal arm residues in solution. In other homeodomain−DNA complexes that have been characterized (*6*), the N-terminal arm makes multiple base pair contacts in the minor groove of substrate DNA. However, **a**1 residues in both the N-terminal arm and at the beginning of H1 (from Lys 0 to Ala 12) were undetected and disordered in the **a**1−α2−DNA ternary structures determined by crystallographic methods (*4*, *42*). Further, it has been recently reported that deletion studies suggest that this arm does not play a significant role in **a**1−α2 binding to DNA (*43*).

The previous experimental data, along with the solution structure data, show that the N-terminal arm is unstructured both free and bound to DNA, and this could account for the relative inability of **a**1 to bind DNA on its own compared to other homeodomains. Favorable enthalpic contacts between N-terminal arm residues and the DNA substrate that might help drive the DNA binding event in other systems are not available for **a**1. Since the **a**1−DNA interaction appears to be so important in vivo, the **a**1 homeodomain must have evolved other mechanisms of DNA binding to compensate for its disadvantageous N-terminal arm.

*The C-Terminus of the DNA Recognition Helix Is Poorly Structured in the MAT**a**1 Structure.* The solution structure and backbone dynamics data indicate that the C-terminal end (residues Arg 53−Lys 57) of the DNA-binding helix (H3) is frayed and largely unstructured. Very low order parameters ($S^2 < 0.3$, Figure 5C), intermediate $^3J_{HNH\alpha}$ coupling constants, and a lack of medium- to long-range NOE connectivities are observed for Ser 56 and Lys 57. Slower backbone motions, in the microsecond to millisecond time range, for residues Lys 52−Lys 57 are inferred from resonance peak line shapes and spectral density model fits that require extra R_{ex} terms. These solution characteristics suggest a poorly folded and perhaps underwound C-terminal turn of the DNA-binding helix. Similar solution behavior at the C-terminal end of the DNA-binding helix has been noted for several homeodomains, including Antennapedia (Antp) (*44*), Oct-1 POU (*40*), and vnd/NK-2 (*45*) homeodomains.

Chemical shift mapping studies (*24*) and NOE data for **a**1 bound to α2, in the absence of DNA (unpublished data from S. M. Baxter and J. S. Anderson), suggest that complete folding of the C-terminal end of the recognition helix must occur upon DNA binding, not as a result of the α2−**a**1 association alone. DNA is required to induce the final bound conformation. Upon comparison of the length of the helix in the free **a**1 structure and the ternary structure of the ternary complex (*4*), DNA binding extends the third helix of **a**1 by three residues. The extension of H3 presumably brings Met 54 and Arg 55 into proper alignment to make contacts in the DNA major groove observed in the **a**1−α2−DNA structure (*4*). It is useful to compare this ordering process with other homeodomains whose structures have been determined in free and DNA-bound states. Antp has eight residues at the end of H3 which become α-helical upon DNA binding; there are 10 in vnd/NK-2 and 5 in en. This suggests that there is less enthalpy released due to hydrogen bond formation in **a**1 compared to other homeodomain systems, as well as less of an entropic penalty due to H3 ordering (*46*, *47*) upon DNA binding. However, the rearrangement of this third helix upon DNA contact alone is apparently not sufficient to drive high-affinity binding. In lieu of additional binding energy, perhaps gained by N-terminal arm interactions in other systems, **a**1 still requires the changes induced by α2 to overcome the energy barriers the native **a**1 protein faces in binding DNA.

Turn 2 Is Not Greatly Affected by α2 Binding. Turn 2 of the homeodomain (T2) is contained in the characteristic DNA binding helix−turn−helix (HTH) structure, and typically, the length and structure of this turn are conserved among homeodomains and HTH binding proteins (*48*). The ^{15}N backbone relaxation data indicate that the DNA-binding HTH portion of the **a**1 homeodomain may not be a rigid DNA-binding unit; rather, it may also have some hinge-like flexibility. Chemical shift mapping studies (*24*) of the free and α2-bound forms of **a**1 do not result in significant chemical shift changes for turn 2 residues upon heterodimerization in the absence of DNA. Any backbone conformational changes observed in this region upon ternary complex formation must occur upon DNA binding. As a result, it is not likely that changes in turn 2 are related to the instructions that α2 gives to induce high-affinity **a**1−DNA binding.

Changes in the Loop 1 Structure Influence Tertiary Structure and DNA Binding Interface. The first turn, a loop connecting H1 and H2, forms part of the hydrophobic patch recognized by the C-terminal tail extension of α2 (*4*, *24*). Amide proton line widths and ^{15}N relaxation data (Figure 5B) suggest that loop 1 residues undergo exchange processes on the microsecond to millisecond time scale; therefore, it is likely that multiple conformations of loop 1 exist in solution. Flexibility in the first loop has also been noted in solution studies of the Oct-1 POU domain (*40*), another homeodomain that requires a protein cofactor, VP16, to function (*49*, *50*).

An important structural role for loop 1 may be the relative orientation of H1 with respect to H2 and H3. The **a**1 structure in solution differs in tertiary arrangement of its helices relative to other homeodomain structures (Table 2). By randomly removing NOEs from the structure calculations, we determine the same alignment of helices, so it is unlikely that a few maverick NOEs in the original data set are forcing H1 to form uncharacteristic angles with respect to H2 and H3 in the ensemble of solution structures. In addition, backbone relaxation behavior of residues at the interface between H1 and H3 is characteristic of motions on the microsecond to millisecond time scale, further supporting movement of the helices relative to each other. As a result, we propose that the flexibility of the loop 1 backbone allows a wider range of hinge motions in the **a**1 homeodomain compared to other structures, resulting in an average **a**1 solution structure with unique interhelical angles.

The **a**1 homeodomain sequence is markedly different from other homeodomains in the loop 1 region (Figure 1). A conserved tyrosine, Tyr 25, found in Antp, en, and α2 homeodomains makes DNA sugar contact (*51*). This residue is replaced with a serine in the **a**1 homeodomain. Ser 25 does not make the analogous polar contact with DNA in the **a**1−α2−DNA complex (*4*). Next, the Antp and en homeodomains have an arginine at position 24; a glutamine is found at this position in **a**1. The side chains of Phe 20, Arg 24, and Tyr 25 form a hydrophobic stack at the C-terminal end of H1 in the Antp and en structures. The shorter side chains found in the **a**1 sequence cannot form this stack, perhaps destabilizing the C-terminal end of H1. We propose that it is this segment of loop 1 sequence that results in the loss of a conserved polar DNA contact, poor side chain stacking, and increased H1−H2 hinge flexibility in the native **a**1 homeodomain. This poorly folded and flexible loop region and the ineffective N-terminal arm combine to make the **a**1 homeodomain a poor DNA-binding protein on its own.

Changes in Loop 1 Hydrogen Bonding and Side Chain Packing Are Associated with Ternary Complex Formation. Previous structural studies (*4*, *40*, *51*) have highlighted the relationship between loop 1 residues and the DNA-binding helix in homeodomains. In particular, the invariant Arg 53 residue makes a network of hydrogen bonds involving bound water molecules, the DNA backbone, and backbone atoms of Leu 26 and Gln 24 in both ternary complex structures (*4*, *42*). In the solution structure of free **a**1, we also observe the Arg 53 side chain in proximity to the backbone atoms of Leu 26, dictated by the NOEs observed between the two residues. The Leu 26 backbone amide exchanges rapidly with solvent (*24*), suggesting that the amide is exposed at the surface and not part of a rigid backbone structure in the free form of the protein. In addition, the ternary complex highlights a hydrogen bond between the NH1 group of Arg

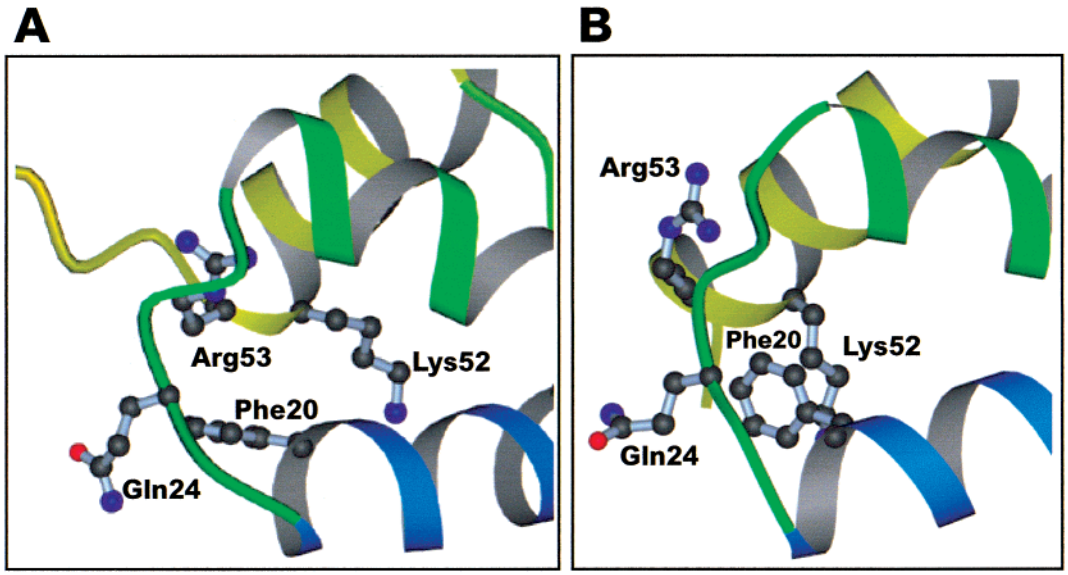

FIGURE 6: Comparison of the loop 1 region of the **a**1 homeodomain, (A) free and (B) bound in the ternary complex (**a**1−α2−*hsg* operator). Limited NOE data defining the side chain atoms of the residues in loop 1 and the ^{15}N backbone dynamics data suggest that the region is not rigidly structured, allowing side chain packing changes. This flexibility allows for van der Waals contacts, based on NOE-derived distance constraints, between Phe 20 and Arg 53 in the **a**1 solution structure. The differences seen in this region, comparing the free and bound **a**1 structures, suggest a mechanism linking the **a**1−α2 protein−protein interaction to changes in the DNA-binding helix of **a**1. The figures were generated from the representative solution structure of **a**1 (A) free in solution and the X-ray coordinates (B) from the **a**1−α2−*hsg* operator complex (1YRN). The Phe 20, Gln 24, Arg, and Lys side chains are depicted in both structures. The figure was created with the program MolScript (*56*).

53 and the carbonyl oxygen of Gln 24. In the free **a**1 structure, we find the backbone carbonyl of Gln 24 and HE of Arg 53 within hydrogen bonding distance, instead of the NH1 protons. This network of hydrogen bonding and side chain contacts serves to anchor H3 to the loop between H1 and H2 in both the free and bound forms, although subtle changes in hydrogen bonding may occur upon ternary complex formation.

Some of the largest differences in backbone torsion angles between the free and ternary complex-bound structures of **a**1 are observed in loop 1 at Gln 24, Ser 25, and Leu 26. Ser 25 is a notably broadened, doubled peak with an unusual ^{15}N frequency of 113.76 ppm in all free **a**1 spectra that sharpens upon heterodimerization in α2-bound **a**1 NMR spectra (*24*). Marked changes were also observed in the amide proton backbone chemical shift, NOE pattern, and coupling constant at Leu 26 upon α2 binding alone. Backbone changes in the segment of residues 24−26 could be effected by the contact observed in the ternary complex (*4*) between the side chain of Asn 27 and the backbone amide of Thr 60 in α2**.** These changes could be the basis of the instructions that the α2 tail gives to **a**1 to convert it from a weak to a strong DNA-binding protein.

The specific change in the loop 1 structure involves the van der Waals stacking of side chains at the C-terminal end of H1. The highly conserved Phe 20 residue is not fully buried in the interior of the homeodomain in either the free **a**1 or the free Oct-1 structures. Aliphatic contacts between Phe 20 and residues in the DNA recognition helix of Oct-1 have been noted previously (*40*)**.** In agreement, we see van der Waals contacts, based on NOE-derived distance constraints, between Phe 20 and Arg 53 in the **a**1 structure (Figure 6A). The aliphatic portion of Arg 53 provides hydrophobic burial for one side of the Phe 20 side chain in the representative **a**1 structure. It is at this point, at Arg 53, that the DNA-binding helix (H3) begins fraying in solution. The long side chain of Arg 24 that buries the Phe 20 side chain in the engrailed structure is replaced with glutamine in both **a**1 and Oct-1. Gln 24 does not make efficient van der Waals contact with the aromatic ring of Phe 20 in the representative **a**1 solution structure (Figure 6A).

Binding of the α2 C-terminal tail results in more efficient side chain packing of the aliphatic portion of the Gln 24 side chain against the aromatic ring of Phe 20 in **a**1 (Figure 6B). In the bound complex (*4*), the aliphatic side chain of Lys 52, instead of Arg 53, stacks against the Phe 20 side chain, as a result of a tighter turn of the α-helix when **a**1 is bound to DNA. Thus, a change at Phe 20 could be transmitted to residues in the DNA recognition helix, effectively providing additional ordering energy for the folding of the C-terminal helix observed upon DNA binding.

*The α2 Interaction Stabilizes the DNA Binding Conformation of the **a**1 Homeodomain.* Since the **a**1 homeodomain does not appear to use adjacent, flanking sequences for protein−protein interactions, it must take another approach to gain combinatorial advantage. Flexibility coupled with protein−protein interactions could combine to modulate DNA binding affinity of homeodomains such as **a**1 and the Oct-1 POU domain. It is true that flexibility in protein structure could lead to an entropic penalty upon DNA binding that might result in lowered binding affinities. But, flexibility can serve as a control mechanism in vivo since these homeodomains are only able to effectively bind DNA under the proper circumstances, when their protein cofactors are present.

Others have noted (*21*, *22*) that most of the specific energy for binding DNA observed upon **a**1−α2 heterodimerization comes from the effect of α2 on the **a**1 structure rather than the effective tethering or recruiting of one homeodomain for another. In particular, Stark and co-workers (*22*) have

recently demonstrated that a peptide with a sequence identical to the C-terminal tail of α2 is sufficient to induce changes in **a**1, making it competent to bind DNA tightly and specifically in vitro. These studies suggest that the α2 peptide−**a**1 homeodomain interaction provides the energetic switch that turns **a**1 into a functional repressor.

We propose that there are at least three ways in which the **a**1 protein becomes more ordered upon heterodimerization and DNA binding. Two involve secondary structure. First is the extension of the DNA recognition helix (H3) by three residues, and second is the rearrangement of residues in loop 1 which influences side chain packing at the end of helix 1. On the basis of backbone displacements, dihedral angle changes, and chemical shift changes, at least four residues in loop 1 (Lys 23−Asn 27) must rearrange upon heterodimerization. The third component of folding involves changes in tertiary structure: the restriction of hinge motions at loop 1 and the reorientation of H1 with respect to H2 and H3. These folding operations are related to each other; in particular, the reorientation of H1 may not be possible without the rearrangement of residues in loop 1.

These ordering changes are triggered by the binding of the α2 tail to the loop 1 region of the **a**1 homeodomain. We propose that this α2−**a**1 interface serves to reduce the flexibility of loop 1 in **a**1. Concomitantly, the changes in the stacking interactions in the loop 1 region cooperatively stabilize the functional DNA-bound structure of the **a**1 recognition helix. The effect of this α2 tail-induced ordering must be a reduction in the entropic barrier the native **a**1 homeodomain faces in folding into its high-affinity DNA-binding form. As a result, the free energy of heterodimer−DNA association is more favorable than that of the **a**1 monomer−DNA association.

These results reinforce the idea of a stepwise series of associations that confer specificity and tight binding on heterodimeric ternary complexes (*52*, *53*). While our studies highlight structural changes that occur in the **a**1 homeodomain structure upon ternary complex formation, we need information from calorimetric studies to characterize the thermodynamic fingerprint and relative energetic importance of the protein−protein and heterodimer−DNA association events. In addition, we plan to study **a**1 homeodomain sequence variants to test the cooperative linkage between side chain interactions in loop 1 and high-affinity DNA binding. Nonetheless, subtle, intricate changes in structure and dynamics hold the key to specificity and affinity in this system and provide a model for other homeodomain proteins that require cofactors to function properly.

ACKNOWLEDGMENT

We thank Jacquelyn Fetrow, Jeff Bell, and Patrick Van Roey for helpful discussions. We thank Lynn McNaughton for assistance with NMR data acquisition. We acknowledge analytical help from the Wadsworth Center Mass Spectroscopy Core. We thank the Wadsworth Center for support and use of the NMR Structural Biology Core.

SUPPORTING INFORMATION AVAILABLE

Table of proton, carbon, and nitrogen NMR resonance assignments. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Gehring, W. J., Affolter, M., and Burglin, T. (1994) *Annu. Rev. Biochem. 63*, 487−526.
2. Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D., and Pabo, C. O. (1991) *Cell 67*, 517−528.
3. Klemm, J. D., Rould, M. A., Aurora, R., Herr, W., and Pabo, C. O. (1994) *Cell 77*, 21−32.
4. Li, T., Stark, M. R., Johnson, A. D., and Wolberger, C. (1995) *Science 270*, 262−269.
5. Wilson, D. S., Guenther, B., Desplan, C., and Kuriyan, J. (1995) *Cell 82*, 709−719.
6. Fraenkel, E., Rould, M. A., Chambers, K. A., and Pabo, C. O. (1998) *J. Mol. Biol. 284*, 351−361.
7. Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K. (1999) *Nature 397*, 714−719.
8. Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L., and Wolberger, C. (1999) *Cell 96*, 587−597.
9. Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J., and Wuthrich, K. (1990) *EMBO J. 9*, 3085−3092.
10. Qian, Y. Q., Otting, G., Billeter, M., Muller, M., Gehring, W., and Wuthrich, K. (1993) *J. Mol. Biol. 234*, 1070−1083.
11. Gruschus, J. M., Tsao, D. H., Wang, L. H., Nirenberg, M., and Ferretti, J. A. (1999) *J. Mol. Biol. 289*, 529−545.
12. Fraenkel, E., and Pabo, C. O. (1998) *Nat. Struct. Biol. 5*, 692−697.
13. Laughon, A. (1991) *Biochemistry 30*, 11357−11367.
14. Keleher, C. A., Goutte, C., and Johnson, A. D. (1988) *Cell 53*, 927−936.
15. Passmore, S., Elble, R., and Tye, B. K. (1989) *Genes Dev. 3*, 921−935.
16. Tan, S., and Richmond, T. J. (1998) *Nature 391*, 660−666.
17. Dranginis, A. M. (1990) *Nature 347*, 682−685.
18. Goutte, C., and Johnson, A. D. (1993) *J. Mol. Biol. 233*, 359−371.
19. Phillips, C. L., Stark, M. R., Johnson, A. D., and Dahlquist, F. W. (1994) *Biochemistry 33*, 9294−9302.
20. Komachi, K., and Johnson, A. D. (1997) *Mol. Cell. Biol. 17*, 6023−6028.
21. Jin, Y., Zhong, H., and Vershon, A. K. (1999) *Mol. Cell. Biol. 19*, 585−593.
22. Stark, M. R., Escher, D., and Johnson, A. D. (1999) *EMBO J. 18*, 1621−1629.
23. Vershon, A. K., Jin, Y., and Johnson, A. D. (1995) *Genes Dev. 9*, 182−192.
24. Baxter, S. M., Gontrum, D. M., Phillips, C. L., Roth, A. F., and Dahlquist, F. W. (1994) *Biochemistry 33*, 15309−15320.
25. Forman, M. D., Stack, R. F., Masters, P. S., Hauer, C. R., and Baxter, S. M. (1998) *Protein Sci. 7*, 500−503.
26. Zwahlen, C., Gardner, K., Sarma, S., Horita, D., Byrd, R., and Kay, L. (1998) *J. Am. Chem. Soc. 120*, 7617−7625.
27. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
28. Stone, M. J., Chandrasekhar, K., Holmgren, A., Wright, P. E., and Dyson, H. J. (1993) *Biochemistry 32*, 426−435.
29. Abragham, A. (1961) *The principles of nuclear magnetism*, Clarendon Press, Oxford, England.
30. Hiyama, Y., Niu, C., Silverton, J. V., Bavoso, A., and Torchia, D. A. (1988) *J. Am. Chem. Soc. 110*, 2378.
31. Mandel, A. M., Akke, M., and Palmer, A. G., III (1995) *J. Mol. Biol. 246*, 144−163.
32. Brunger, A. (1992) *Xplor*, version 3.851, Yale University Press, New Haven, CT.
33. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
34. Kelley, L. A., Gardner, S. P., and Sutcliffe, M. J. (1996) *Protein Eng. 9*, 1063−1065.
35. Nilges, M., and O'Donoghue, S. (1998) *Prog. NMR Spectrosc. 32*, 107−139.
36. Kabsch, W., and Sander, C. (1983) *FEBS Lett. 155*, 179−182.
37. Doreleijers, J. F., Raves, M. L., Rullmann, T., and Kaptein, R. (1999) *J. Biomol. NMR 14*, 123−132.

38. Christopher, J. A., Swanson, R., and Baldwin, T. O. (1996) *Comput. Chem. Eng. 20*, 339−345.
39. Aurora, R., and Rose, G. D. (1998) *Protein Sci. 7*, 21−38.
40. Cox, M., van Tilborg, P. J., de Laat, W., Boelens, R., van Leeuwen, H. C., van der Vliet, P. C., and Kaptein, R. (1995) *J. Biomol. NMR 6*, 23−32.
41. Ippel, H., Larsson, G., Behravan, G., Zdunek, J., Lundqvist, M., Schleucher, J., Lycksell, P. O., and Wijmenga, S. (1999) *J. Mol. Biol. 288*, 689−703.
42. Li, T., Jin, Y., Vershon, A. K., and Wolberger, C. (1998) *Nucleic Acids Res. 26*, 5707−5718.
43. Wolberger, C. (1999) *Annu. Rev. Biophys. Biomol. Struct. 28*, 29−56.
44. Billeter, M., Qian, Y., Otting, G., Muller, M., Gehring, W. J., and Wuthrich, K. (1990) *J. Mol. Biol. 214*, 183−197.
45. Tsao, D. H., Gruschus, J. M., Wang, L. H., Nirenberg, M., and Ferretti, J. A. (1995) *J. Mol. Biol. 251*, 297−307.
46. Spolar, R. S., and Record, M. T., Jr. (1994) *Science 263*, 777−784.
47. Bagby, S., Arrowsmith, C. H., and Ikura, M. (1998) *Biochem. Cell Biol. 76*, 368−378.
48. Tan, S., and Richmond, T. J. (1998) *Curr. Opin. Struct. Biol. 8*, 41−48.
49. Lai, J. S., Cleary, M. A., and Herr, W. (1992) *Genes Dev. 6*, 2058−2065.
50. Pomerantz, J. L., Kristie, T. M., and Sharp, P. A. (1992) *Genes Dev. 6*, 2047−2057.
51. Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G., and Wuthrich, K. (1994) *Cell 78*, 211−223.
52. Sanchez, M., Jennings, P. A., and Murre, C. (1997) *Mol. Cell. Biol. 17*, 5369−5376.
53. Jabet, C., Gitti, R., Summers, M. F., and Wolberger, C. (1999) *J. Mol. Biol. 291*, 521−530.
54. Ferrin, T., Huang, C., Jarvis, L., and Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
55. Ferrin, T., Huang, C., Jarvis, L., and Langridge, R. (1988) *J. Mol. Graphics 6*, 36−37.
56. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI000677Z