Structural Studies of the High Mobility Group Globular Domain and Basic Tail of HMG-D Bound to Disulfide Cross-Linked DNA†

Linda K. Dow,‡,§ David N. M. Jones,*,|| Scot A. Wolfe,⊥ Gregory L. Verdine,# and Mair E. A. Churchill*,||

Department of Pharmacology, University of Colorado Health Sciences Center, C236, 4200 East Ninth Avenue, Denver, Colorado 80262, Center for Biophysics and Computational Biology, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Received March 30, 2000; Revised Manuscript Received June 8, 2000

ABSTRACT: HMG-D is an abundant high mobility group chromosomal protein present during early embryogenesis in *Drosophila melanogaster*. It is a non-sequence-specific member of a protein family that uses the HMG domain for binding to DNA in the minor groove. The highly charged C-terminal tail of HMG-D contains AK motifs that contribute to high-affinity non-sequence-specific DNA binding. To understand the interactions of the HMG domain and C-terminal tail of HMG-D with DNA in solution, a complex between a high-affinity truncated form of the protein and a disulfide cross-linked DNA fragment was studied using heteronuclear NMR techniques. Despite its relatively high affinity for the single "prebent" site on the DNA, $K_d = 1.4$ nM, HMG-D forms a non-sequence-specific complex with the DNA as indicated by exchange broadening of the protein resonances at the DNA interface in solution. The secondary structural elements of the protein are preserved when the protein is complexed with the DNA, and the DNA-binding interface maps to the regions of the protein where the largest chemical shift differences occur. The C-terminal tail of HMG-D confers high-affinity DNA binding, has an undefined structure, and appears to make direct contacts in the major groove of DNA via residues that are potentially regulated by phosphorylation. We conclude that while the HMG domain of HMG-D recognizes DNA with a mode of binding similar to that used by the sequence-specific HMG domain transcription factors, there are noteworthy differences in the structure and interactions of the C-terminal end of the DNA-binding domain and the C-terminal tail.

The nonhistone high mobility group (HMG)[1] proteins that are abundant in higher eukaryotes have roles in modulating the structure and activity of chromatin (reviewed in refs *1* and *2*) and in activating specific processes such as recombination (*3*) and transcription (*4*). The archetypes of the HMG1/2 chromosomal protein family are the vertebrate HMG1 and HMG2 proteins which bind to DNA non-sequence-specifically using two copies of a conserved 75-residue DNA-binding motif known as the HMG box or HMG domain (*1, 5*). Other chromosomal proteins that have similar binding properties include the yeast proteins NHP6A and NHP6B (*6*), as well as the *Drosophila melanogaster* chromosomal protein HMG-D. These latter proteins, however, contain a single HMG domain and N- or C-terminal tails composed of basic and/or acidic regions (*7*−*9*). A single HMG domain is also found in a number of sequence-specific transcription factors, such as lymphoid enhancer factor, LEF-1 (*10*), and the mammalian testis determining factor, SRY (*11*). Regardless of domain number or specificity, HMG domains preferentially bind to DNA that is prebent, such as cisplatin-modified DNA, disulfide cross-linked DNA, or easily deformable DNA sequences (*9*, *12*). This characteristic bending, which defines a mode of DNA binding known as architecture specificity, may bring neighboring bound proteins into closer proximity to facilitate transcriptional activation and the compaction of chromosomal DNA (*13*).

Solution structural studies using NMR have revealed that the uncomplexed chromosomal HMG domains adopt a stably folded L-shaped structure composed of three α-helices (*14*−*17*). In contrast, uncomplexed sequence-specific proteins, such as SOX-4, appear to require DNA binding in order to achieve a stable fold (*18*). In either case, the mode by which the HMG domain proteins bind to DNA is similar and has been revealed by the NMR structures of sequence-specific HMG domains from LEF-1 and SRY complexed with DNA (*19*−*21*), the crystal structures of the HMG domains of

† This work was supported by the American Heart Association, the American Cancer Society (RPG-97-060-01), NIH Grant RO1-GM59456 to M.E.A.C., and NIH Molecular Biophysics predoctoral training grant, University of Illinois Graduate College Travel Grant, and Texaco Foundation Scholarship to L.K.D.

* To whom correspondence should be addressed: (303) 315-0427 (telephone); (303) 315-7097 (fax); mair.churchill@uchsc.edu or david.jones@uchsc.edu (e-mail).

‡ University of Illinois at Urbana−Champaign.

§ Current address: Eli Lilly and Co., Indianapolis, IN.

|| University of Colorado Health Sciences Center.

⊥ Massachusetts Institute of Technology.

Harvard University.

[1] Abbreviations: DNA, deoxyribonucleic acid; DTT, dithiothreitol; cisplatin, *cis*-diamminedichloroplatinum(II); fid, free induction decay; HMG, high mobility group; IPTG, isopropyl β-D-thiogalactoside; LEF-1, lymphoid enhancer factor −1; NHP6, nonhistone protein 6; NOE, nuclear Overhauser effect; SRY, testis determining factor; TRIS, tris(hydroxymethyl)aminomethane.

10.1021/bi000723v CCC: $19.00
Published on Web 07/14/2000

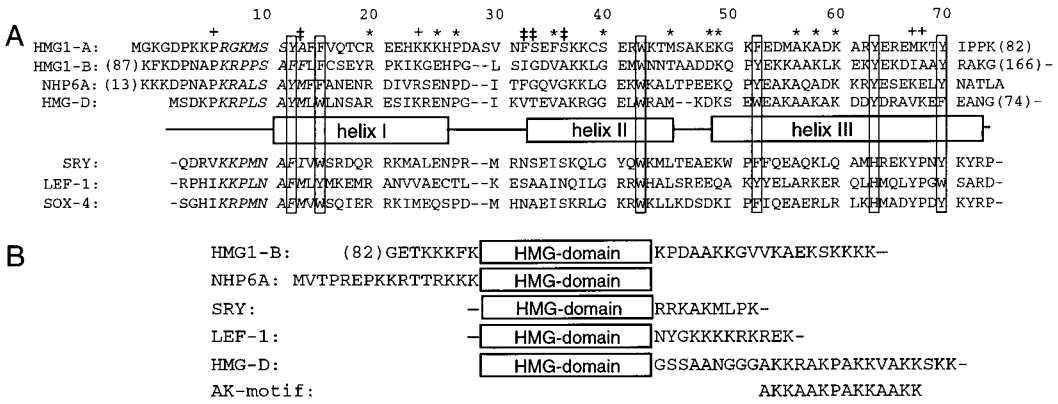

FIGURE 1: HMG domain structure and sequence analysis. (A) HMG domain protein sequences, HMG1 boxes A and B, NHP6A, HMG-D, SRY, LEF-1, and Sox4, are grouped according to their DNA binding specificity and aligned on the basis of HMG-D numbering. Highly conserved residues in italics precede the first DNA intercalating residue, and those outlined in black are conserved hydrophobic core residues. Key: ‡, DNA intercalating residue; *, additional highly conserved residues; +, residues that differ significantly between the sequence-specific and non-sequence-specific HMG domain proteins. The horizontal rectangles represent helices, I, II, and III. (B) Alignment of HMG domain basic tail regions for HMG1 box B, NHP6A, SRY, LEF-1, and HMG-D. The tail motif of HMG-D is aligned with the canonical histone H1 AK motif (*64*).

HMG1 and HMG-D complexed with DNA (*22*, *23*), and numerous biochemical experiments. These studies have shown that the domain binds to a highly distorted minor groove using predominantly hydrophobic contacts with relatively few base-specific hydrogen bonds. Partial DNA intercalation by either a methionine or isoleucine at position 13 in helix I and residues at positions 32 and 33 appear to contribute to DNA kinking and shape complementarity (*23*). The structure of the HMG domain is not substantially perturbed by binding to linear DNA. However, in the structure-specific complex of HMG1 bound to a "prebent" cisplatin-modified DNA, the HMG domain changes more upon DNA binding than does the DNA substrate (*22*).

The affinity and architectural specificity of several HMG domain proteins also depend on regions that lie outside the HMG domain (Figure 1). Basic tails in the sequence-specific proteins lie directly adjacent to the C-terminus of the HMG domain, beginning with the amino acids KYX, where X is any basic residue. The absence of the C-terminal basic tail in LEF-1 decreases its affinity and the degree of DNA bending (*24*). In the yeast chromosomal protein, NHP6A, a basic tail region is found at the N-terminus, residues 2−16; the deletion of this basic tail drastically reduces the ability of NHP6A to bind DNA although it can still bend DNA (*25*).

The "basic tail" region of HMG-D (residues 80−100) begins several residues after the C-terminus of the HMG domain (residues 1−74). The HMG-D basic tail is longer than the HMG1/2 basic regions and is followed by a C-terminal acidic stretch (residues 101−112) like that seen in other HMG1/2 proteins. Deletion of the C-terminal tail of HMG-D results in 2 orders of magnitude lower binding affinity, suggesting that the tail of HMG-D is important for high-affinity DNA binding (*26*, *27*). Interestingly, the HMG domain alone (HMG-D-74), the HMG domain with the basic tail (HMG-D-100), and the full-length protein (HMG-D-112) form DNA circles of the same size, only 55 base pairs, in ligase-mediated circularization assays, indicating that all forms of HMG-D that contain the HMG domain bend and unwind DNA to the same degree (*27*). The HMG-D basic tail is composed of a tandemly repeated "AK" motif that is found in only a few other proteins, most notably the C-terminal tails of vertebrate histone H1 (Figure 1B) (*28*). These tails appear to be unstructured in the absence of DNA (*29*) but may adopt helical structure upon DNA binding as shown by circular dichroism studies with histones and fiber diffraction studies with model peptides (*30*−*32*). Like the C-terminal tail of histone H1, the basic tail of HMG-D contains sequences that may be posttranslationally phosphorylated, possibly modulating the function of the tail (*33*, *34*).

To understand how the tail of HMG-D interacts with DNA and determine what role it may play in sequence specificity, NMR structural studies were conducted on a sample of 60% ^{2}H, 100% ^{13}C, ^{15}N-labeled HMG-D-100 (residues 1−100) complexed with disulfide cross-linked DNA. Secondary structure analyses, NOE contacts, and structure calculations were used to characterize the structure of the HMG domain and the basic tail of HMG-D while bound to DNA. Chemical shift differences between the complexed and free forms of HMG-D provided insight into the regions of the protein affected by DNA binding. The AK motif exerts its affinity-enhancing effects without forming any regular secondary structure. Finally, the interactions of the C-terminal tail of HMG-D-100, as revealed by filtered NOE experiments, suggest that the tail interacts with the DNA major groove in a manner that differs from other HMG domain proteins, thus explaining functional differences.

MATERIALS AND METHODS

Preparation of ^{2}H, ^{13}C, ^{15}N-Labeled HMG-D-100. The preparation of the T7 expression system vector pET13a expressing HMG-D-100 has been described previously (*9*, *35*). Bacterial growth was in M9 minimal media supplemented with 0.01 g/L $FeSO_4 \cdot 7H_2O$, 0.01 g/L thiamin, 10 mL/L basal medium Eagle vitamin solution (Gibco-BRL), and 20 μg/mL kanamycin. [^{15}N]NH_4Cl and [^{13}C]glucose were added at 0.5 and 2 g/L, respectively. Approximately 60% deuterium incorporation was achieved by using 600 mL of D_2O/L of media. All isotopes were obtained from Isotec. Medium was prepared and sterile filtered immediately prior to use. To adapt the bacteria to the deuterium-containing media for optimal protein expression, the following procedure was used (Kevin Gardner, personal communication). Growth was initiated by inoculating 3 mL of LB media (with

kanamycin) with a fresh glycerol stock of *Escherichia coli* strain BL-21 transformed with the pET13-HMG-D-100 plasmid. The culture grew for 4 h at 37 °C with shaking. The cells were pelleted and resuspended with minimal media. The suspension was transferred to 25 mL of supplemented minimal media and grown at 37 °C and 200 rpm until an OD_{590} of 0.5 was reached. The cells were spun at 4 °C, 10K rpm for 15 min, and the cell pellet was resuspended in fresh media. A sufficient amount of the suspension was added to 75 mL of media to yield a starting OD_{590} of approximately 0.1. This culture was grown until it reached an OD_{590} of 0.46. The cells were again pelleted and resuspended in fresh media. A sufficient amount was transferred to 400 mL of media to obtain an OD_{590} of 0.1. The culture was grown as before to an OD_{590} of ~0.5. The resuspension procedure was repeated, transferring the cells to four 2 L flasks each containing 1 L of fresh media. The cultures were grown to an OD_{590} of ~0.5 and were induced with IPTG (1 mM) and grown for 4 h. Cells were centrifuged, and the protein purification was carried out as described previously (*36*). The protein was lyophilized after the final purification step.

Preparation of Cross-Linked DNA. Synthesis of the cross-linked DNA proceeded as previously described with a few exceptions (*9*, *37*). Single-stranded DNA (5′-CGAATCXATTGATTCG-3′) was synthesized on a 394 Applied Biosystems synthesizer at the DNA Synthesis Laboratory at the University of Wisconsin Biotech Center in Madison, WI. X indicates the use of the special phosphoramidite, the site for initiating the cross-linking reaction (*9*, *37*). PAC (5′-dimethoxytrityl-*N*-phenoxyacetyl) phosphoramidites (Pharmacia) were used for the remainder of the synthesis so that a mild deprotection scheme could be used. Fifteen 1 μmol syntheses were completed. The resulting product was lyophilized.

The double-stranded DNA product was separated from single-stranded starting materials using a Mono Q HR 5/5 column (Pharmacia) on a BioCAD SPRINT (PerSeptive Biosystems Inc.) HPLC system. Buffer A contained 10 mM NaOH and 0.5 M NaCl, pH 12, while buffer B contained 10 mM NaOH and 0.9 M NaCl, pH 12. Single-stranded DNA eluted with a mobile phase composition of 24% buffer B in 14 min while double-stranded DNA eluted at 45% buffer B over 10 min. Fractions containing double-stranded DNA were pooled and desalted by exchanging against water using a YM3 membrane (Amicon). The desalted DNA was lyophilized. A final yield of 25.85 mg of DNA was obtained.

To confirm the presence of the cross-link, a sample of DNA was reduced and analyzed on a 20% denaturing polyacrylamide gel. The sample was reduced by incubating approximately 4 μg of DNA and 5 μL of 90% formamide dye with 5 μL of a 100 mM DTT/10 mM TRIS solution (pH 9.0) at 90 °C for 5 min. Samples of nonreduced DNA and single-stranded DNA were run as controls. The DNA migrated as single stranded only after being exposed to strong reducing conditions. The DNA was shown to be homogeneous using electrospray mass spectrometry analysis, which showed a single major peak at 9907 ± 3.4 mass units.

Formation of the ^{2}H, ^{13}C, ^{15}N-Labeled HMG-D-100−Cross-Linked DNA Complex. An NMR sample of cross-linked DNA was prepared using 11.25 mg of DNA in 50 mM sodium acetate, pH 5.0, 10% D_2O, and 0.02% NaN_3 for a final volume of 300 μL. To obtain a 0.9:1 protein:DNA sample, 88.4 μL of stock protein was required. This volume was added to the DNA in three aliquots, a 1 h incubation period being allowed between each addition. The titration was followed by observing the DNA and tryptophan region of a 1D ^{1}H-NOESY spectrum. Water suppression was achieved by replacing the last 90° pulse with a 1-1 pulse (*38*), and States/TPPI was used for quadrature detection in the indirect dimension (*39*). Data were processed with VNMR software (Varian). Spectra were collected at 15, 20, 25, 27, 30, and 37 °C. The final concentration of protein:DNA was approximately 1.50 mM:1.76 mM in 20 mM NaCl, 50 mM sodium acetate, 0.02% NaN_3, and 10% D_2O at pH 5.72.

NMR Spectroscopy. NMR experiments for structure elucidation were conducted on a Varian Unity Inova 600 MHz spectrometer with a Varian 5 mm $^{1}H\{^{13}C,^{15}N\}$ triple axis gradient probe. All experiments were recorded at 25 °C, and States/TPPI was used for quadrature detection in the indirect dimension (*39*). Table 1 lists the experiments used for the backbone, side chain, and NOE assignments of 60% ^{2}H, 100% ^{13}C, ^{15}N-labeled HMG-D-100 in complex with cross-linked DNA. The transmitter was typically positioned at the center of the ^{13}C frequency being detected while radio frequency pulses were applied as phase modulated pulses positioned at the second appropriate carbon frequency. The one exception was the HN(COCA)CB experiment where the transmitter position alternated between the carbonyl and α-carbon frequency ranges.

Experiments were based on those cited in Table 1, with several modifications. Sensitivity enhancement was not used, and the HCCCONH-TOCSY (C), HCCCONH-TOCSY (H), CBCANH, HCCH-TOCSY, HN(CA)CO, ^{13}C NOESY-HSQC, and ^{13}C TOCSY-HSQC all required the addition of deuterium decoupling. Shaped carbon pulses were substituted for hard pulses in many of the experiments to enhance selectivity. Carbon 90° pulses were applied as either E-BURP2 (*40*) or Q5 pulses (*41*) while 180° inversion and refocusing pulses were applied as I-SNOB5 (*42*) and Re-BURP pulses (*40*), respectively. Simultaneous inversion and refocusing were achieved using a hyperbolic secant pulse (*43*, *44*) applied in the middle of the ^{13}C spectrum. A compensating hyperbolic secant pulse was also applied to correct for phase errors (*45*).

Water suppression was achieved by returning the magnetization to the *Z* axis prior to detection by application of Gaussian-shaped selective pulses. Magnetization remaining in the *xy* plane was dephased by the use of gradients applied along the *Z* axis. Gradients were typically applied for 500 μs at 10.9 G/cm.

Pulses in each experiment were individually optimized: E-Burp and Q5 for maximum signal, Gaussian-shaped pulses for water suppression, and gradients for both water suppression and optimal signal.

Hard pulses were applied at field strengths of 34.5 kHz for protons, 19.2 kHz for carbon, and 5.68 kHz for nitrogen. Isotropic mixing in the TOCSY experiments was achieved using the DIPSI2 (*46*) and DIPSI3 (*46*) sequences at spin-lock field strengths of 4.5 and 5.8 kHz, respectively. WALTZ16 (*47*) was used for heteronuclear broad-band decoupling during carbon evolution while Garp (*48*) decoupling was employed for ^{13}C and ^{15}N decoupling during acquisition as needed. One exception was the HNHA

Table 1: NMR Experiments Used in the Structure Determination of an HMG-D-100−DNA Complex

	acquired data (complex points)			sweep width (Hz)					
	t_3	t_2	t_1	w_3	w_2	w_1	mix (ms)	time (h)	ref
backbone assignments									
HSQC		1024 (H)	128 (N)		9000	3000		17	*66*
HN(CO)CA	1024 (H)	32 (N)	42 (C)	9000	3000	4000		160	*67*, *68*
HNCA	1024 (H)	41 (N)	48 (C)	9000	3000	4000		111	*68*
	1024 (H)	34 (N)	56 (C)	8000	2000	5500		112	
HNCO	1024 (H)	38 (N)	48 (C)	9000	3000	4000		51	*68*, *69*
HN(CA)CO	1024 (H)	38 (N)	48 (C)	8000	3000	4000		108	*70*
HN(COCA)CB	1024 (H)	32 (N)	74 (C)	7000	2000	9400		134	*71*
HN(CA)CB	1024 (H)	32 (N)	76 (C)	8000	3000	9500		144	*71*, *72*
CBCANH	1024 (H)	34 (N)	53 (C)	8000	3000	9400		94	*71*
HBHA()NH	1024 (H)	34 (N)	64 (H)	8000	3000	8000		122	*73*
HCA(CO)NH	1024 (H)	32 (N)	48 (H)	8000	3000	3000		92	*73*, *74*
HNHA	1024 (H)	62 (H)	32 (N)	8000	8000	3000		112	*75*, *76*
HCANH	1024 (H)	32 (N)	50 (H)	8000	3000	3000		94	*77*
side chain assignments									
^{15}N TOCSY-HSQC	1024 (H)	34 (N)	96 (H)	8600	3000	8600	120	94	*78*
^{13}C TOCSY-HSQC	1024 (H)	64 (C)	70 (H)	8000	6600	8000	55	64	*78*
HCCH-TOCSY	1024 (H)	66 (C)	60 (H)	8000	10000	8000	30	114	*67*, *79*
HCCCONH-TOCSY (C)	1024 (H)	32 (N)	72 (C)	8000	3000	9400	27	137	*80*−*82*
HCCCONH-TOCSY (H)	1024 (H)	32 (N)	98 (H)	8000	2000	8000	27	134	*80*−*82*
through-space interactions									
^{15}N NOESY-HSQC	1024 (H)	32 (N)	90 (H)	9000	3000	9000	250	94	*83*
^{15}N NOESY-HSQC	1024 (H)	34 (N)	90 (H)	9000	3000	9000	150	94	*83*
^{15}N NOESY-HSQC (Lys/Arg optimized)	1024 (H)	40 (N)	80 (H)	8000	4020	8000	150	98	*83*
^{15}N HSQC-NOESY-HSQC	1024 (H)	40 (N)	40 (N)	8000	3000	3000	150	74	*84*
^{13}C NOESY-HSQC	1024 (H)	72 (C)	125 (H)	8000	6600	8000	150	139	*83*
^{12}C-filtered NOESY-HSQC	1024 (H)	32 (N)	90 (H)	8000	2000	10000	150	178	*85*

Table 2: Structure Calculation Summary[a]

	residues used from		secondary structure analyzed	no. of structures superimposed	RMSD (Å)
	complexed protein (no. of NOEs)	free protein (no. of NOEs)			
hybrid structures					
core	11−48 (202)	1−10, 49−74 (216)	H1 (18−25)	28	0.72
			H2 (33−38)	18	0.52
termini	1−10, 49−74 (178)	11−48 (304)	H3 (51−59)	15	0.89
			H3 (59−72)	24	1.20
tail	75−100 (71)	1−74 (763)	NA	NA	NA
full complex					
domain	1−74 (382)		H1 (18−25)	17	0.74
			H2 (33−38)	11	0.51
			H3 (51−59)	9	0.74
			H3 (59−72)	17	1.10

[a] Abbreviations: NA, not applicable; H1, helix I; H2, helix II; H3, helix III.

experiment which used WALTZ16 during acquisition. Field strengths for WALTZ16, ^{13}C Garp, and ^{15}N Garp decoupling schemes were 4.5, 4.0, and 1.1 kHz, respectively. Selective decoupling of the carbonyl or α-carbon region was achieved using either a hyperbolic secant pulse (*43*, *44*) or a seduce profile (*49*) modulated with an MLEV4 (*50*) supercycle. Deuterium decoupling was achieved with a pulse applied at a field strength of 1.0 kHz.

Data were processed and analyzed with Felix97 software (MSI Inc.) on a Silicon Graphics R5000 O2 workstation. Prior to transformation, data were typically multiplied by a square sine-bell function and zero-filled to 128 and 256 points in the indirect dimensions. The first points of each fid were linear predicted to prevent baseline distortion. Chemical shifts were referenced to external DSS (*51*), using the frequency ratios presented by Wishart et al. (*52*).

Structure Calculations. Distance constraints were predominantly derived from ^{15}N-NOESY-HSQC and ^{15}N-HSQC-NOESY-HSQC spectra. The ^{13}C-NOESY-HSQC experiment provided additional NOEs for the interactions of the aromatic rings. A total of 453 NOEs were classified as being either strong (<2.9 Å), medium (<3.5 Å), or weak (<5.0 Å) on the basis of their relative intensities to NH−NH helical interactions.

Four separate distance constraint lists were assembled from this basis set as summarized in Table 2. The first list included 382 NOEs from residues 1−74 of the complexed protein. No additional constraints for dihedral angles or hydrogen bonds were employed. The second constraint list was designed to study the tail alone. A total of 71 NOEs from the last 26 residues of the complexed protein were appended to the complete NOE constraint list derived for the structure of the free form of HMG-D-74 (*14*). Additional constraint information for dihedral angle and hydrogen bond data from the free protein was also included. The third and fourth constraint lists were hybrid lists combining regions of the

complex structure with regions of the free protein to improve the stability of the overall structure in the calculations. Dihedral angle and hydrogen bond constraints were only included if they spanned residues that were completely contained within the boundaries of the free protein region of the hybrid structure. NOE constraints from the free protein that were not NH−H or aromatic ring NOEs were not used. This eliminated 243 NOEs that were used in the structure determination of the free protein. A set of 30 structures were calculated for the tail construct while 50 structures were calculated for each of the remaining three cases.

Structures were calculated using a simulated annealing protocol implemented in the program X-PLOR (*53*), beginning from a structure with randomized ϕ and ψ angles. NOE constraints involving groups with equivalent or nonstereospecifically assigned protons were treated with R^{-6} summation (*54*). No corrections were added for multiplicities or degeneracies (*55*). Two methods were then employed to assess the quality of the resulting structures. The first criterion was based on the NOE energy term. Those structures with unusually high NOE energies were removed from further analysis. The second technique screened for structures that possessed low NOE energy terms but assumed random extended conformations. These latter cases were identified by superimposing and analyzing the ensemble of structures with the program Superpose (*56*), specialized X-PLOR scripts (*57*), and the visualization program VMD (*58*). The script aligned the calculated structures to a representative structure and computed an average structure. This average structure was minimized to establish realistic bond lengths and angles. Each structure was then realigned to the minimized average, and a new average structure was obtained and minimized. The process was repeated until the average structures converged. An RMSD was then calculated between each member and the average. Those that deviated significantly from the average ($>\sim$8 Å) were removed from further analysis. The process of calculating an average was then repeated with the reduced ensemble in order to obtain the most accurate representative average structure. Finally, each structure was evaluated relative to this final, average, minimized structure.

RESULTS

Protein Resonance Assignments. NMR was used to analyze the interactions between a disulfide cross-linked DNA fragment and the HMG-D construct for which it had the highest affinity, HMG-D-100. The NMR sample contained a 0.9:1 ratio of protein:DNA as judged by electrophoretic mobility shift assays (Figure 2) and 1D proton NMR analyses. The optimal sample temperature of 25 °C was determined by assessing the quality of 1D ^{1}H-NOESY spectra collected at 15, 20, 25, 27, 30, and 37 °C. The NMR experiments used in this work are listed in Table 1.

The resonance assignments are summarized in a labeled ^{15}N-HSQC spectrum shown in Figure 3. Only 9% of the backbone chemical shifts could not be assigned, with the vast majority of these missing assignments falling within the HMG domain itself. The chemical shifts for the main chain atoms of residues 75−100 in the basic tail were easily assigned with the sole exception of the backbone amide resonance of arginine 87. A majority of the side chain resonances were also assigned, although the degree of degeneracy associated with the abundance of lysine residues prevented unequivocal assignment of many side chain protons and contributed to uncertainty in making β, γ, and δ proton designations.

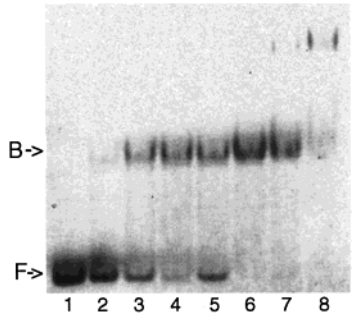

FIGURE 2: HMG-D binding to disulfide cross-linked DNA. The titration of HMG-D-100 protein into DNA, at concentrations comparable to those used in the NMR experiments, was analyzed by native polyacrylamide gel electrophoresis. Ratios of protein:DNA varied from 0.5:1 to 2:1, lanes 2−8. A 1:1 complex was formed as expected, with higher order complexes forming as the ratio increased to 2:1.

The Global Fold of the HMG Domain Is Conserved on Binding to DNA. An initial assessment of the secondary structure present in the complex of HMG-D bound to DNA was obtained using the chemical shift index (CSI) program (*52*). Since the protein was deuterated, the Cα random coil chemical shifts were adjusted upfield by 0.29 ppm (0.39 ppm for glycine residues) to account for the one-bond deuterium isotope effect (*59*). The effect that deuterium has on the α-carbon when it is two or three bonds removed is less predictable due to the dependence on the probability of deuterium being present at the two- and three-bond positions and the dependence on the incorporation of deuterium during the biosynthesis of the residue. When the Cα random coil shifts were adjusted for the full isotope effect of −0.50 ppm, the consensus shift analysis results were the same as with the one-bond adjustment, indicating that the dominating contributor to the Cα shift appears to be due to the one-bond effect. Therefore, the correction for only the one-bond effect was used.

The deviation of chemical shifts from random coil values for Cα, CO, and Hα and the consensus chemical shift index are shown in Figure 4. The values obtained for the HMG domain indicate α-helical structure for residues 11−13, 18−25, 32−38, and 51−72. The absence of assignments for some residues causes apparent helix breaks in helices I and II, as discussed below. Helix I may be complete in the structure of the complex, but on the basis of chemical shift, it terminates two residues prior to that in the structure of the free protein. Helix II appears to begin one residue earlier in the structure of the complex compared to the free protein and stops after residue 38 prior to the Gly-Gly sequence. Residues 41−43 remain unassigned, but the shifts for residue 44 are representative of helical content, indicating that helical structure is most likely maintained throughout the unobserved region. Helix III appears to be well formed between residues 51−72 and is similar to helix III as seen in the structure of the free protein, residues 49−73 (*14*). Despite the unknown effect of the proximity of the DNA on chemical shifts, the CSI analysis appropriately predicts the three helical regions

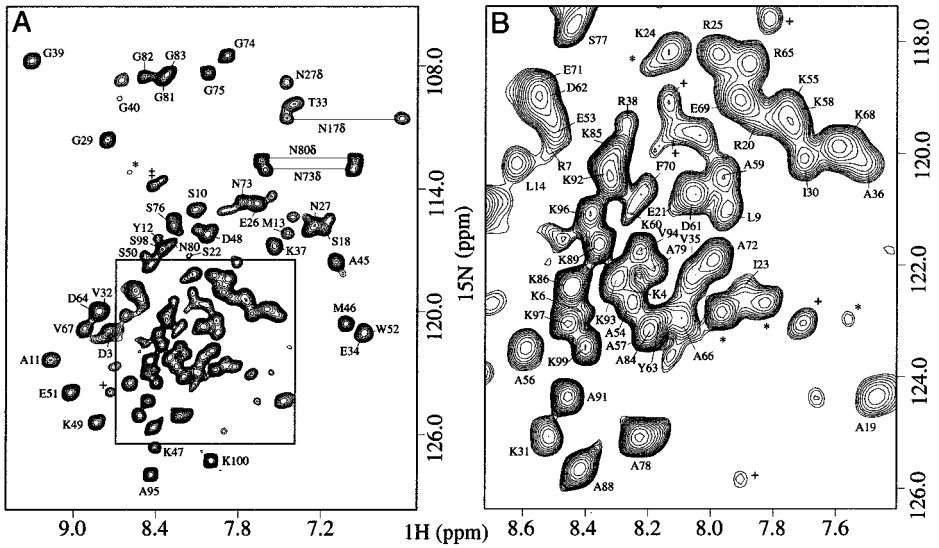

FIGURE 3: Resonance assignments. (A) Identified ^{15}N-HSQC spectrum of HMG-D-100 complexed with cross-linked DNA, with the boxed region expanded in (B). Each peak corresponds to the amide proton and nitrogen shift of a given spin system. Two peaks, marked with a ‡, have anomalous $i-1$ α- and β-carbon chemical shifts, lack NH−H NOEs, and cannot be assigned. The peaks marked with an ∗ remain unassigned. An alanine occurs at the $i-1$ position for each of these peaks, and the i β-carbon resonances are representative of shifts typically seen for asparagine, aspartic acid, isoleucine, leucine, phenylalanine, or tyrosine. Such sequences occur only three times in the protein and have already been clearly assigned. The paucity of NOEs suggests that these residues do not originate from a stable conformation of the globular region of the protein but may originate from the unstructured tail, Ala79−Asn80. Of the remaining unassigned peaks in the HSQC spectrum, many are not found in the backbone experiments or are weak and diffuse peaks with incomplete spin systems and few NH−H NOEs. It is likely that at least a portion of the missing backbone assignments can be ascribed to these peaks, but insufficient information precludes their placement within the protein.

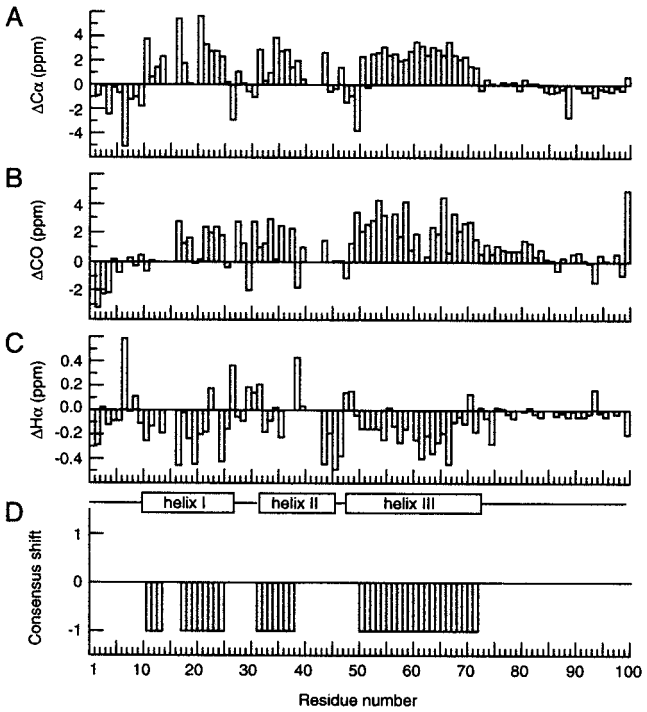

FIGURE 4: Chemical shift index analysis of HMG-D-100 complexed with DNA. Plots A, B, and C show the chemical shift deviations from random coil for Cα, CO, and Hα, respectively. Plot D shows the consensus chemical shift index as determined with the CSI program (*86*). The indicated helices demonstrate the position of α-helices in the structure of the free protein. Chemical shift assignments are missing for residues 15, 16, and 41−43.

observed previously in the structures of HMG-D free and bound to linear DNA.

The CSI results were confirmed by NOE analysis and structure calculations. The distance constraints used in the structure calculations were derived from a total of 453 NOEs. These NOEs fully support the helical structure assignments discussed above (Figure 5A). Although only 19 long-range ($>i,i+4$) interactions were seen for the complexed protein compared to 76 in the free protein (*14*), these NOEs confirm that the bound protein has maintained a fold similar to that of the free protein and are consistent with the positions of the residues of the protein in the HMG-D−DNA crystal, as seen in Figure 5B (*23*). In particular, NOE data linking the amide proton of Lys4 to the ring of Phe70, as well as data supporting an interaction between the side chains of Pro8 and Tyr63, indicate that the N- and C-termini of the protein align similarly in the free and bound forms of the protein. NOEs between ring protons of Trp15 and Trp52 along with NOEs between Trp43 and Tyr12 show that contacts in both hydrophobic cores are also maintained.

Structure calculations of the complex were carried out using several complementary approaches, summarized in Table 2. Constraints from isolated regions of the free and bound forms of the protein were used to generate hybrid constraint lists, so that independent analyses of the basic tail and each of the three α-helices could be performed. The first approach used constraints from the HMG domain of the complexed protein to examine whether structural changes had occurred with DNA binding. The second approach was to study the tail from the complex, while it was positioned at the end of the well-characterized globular domain of the free protein (*14*), so that it could be analyzed independently of the less well determined regions of the HMG domain of the complex. The third and fourth approaches used chimeras of the long and short wing of the free and complexed proteins to improve the stability of the overall structure in the calculations (*60*).

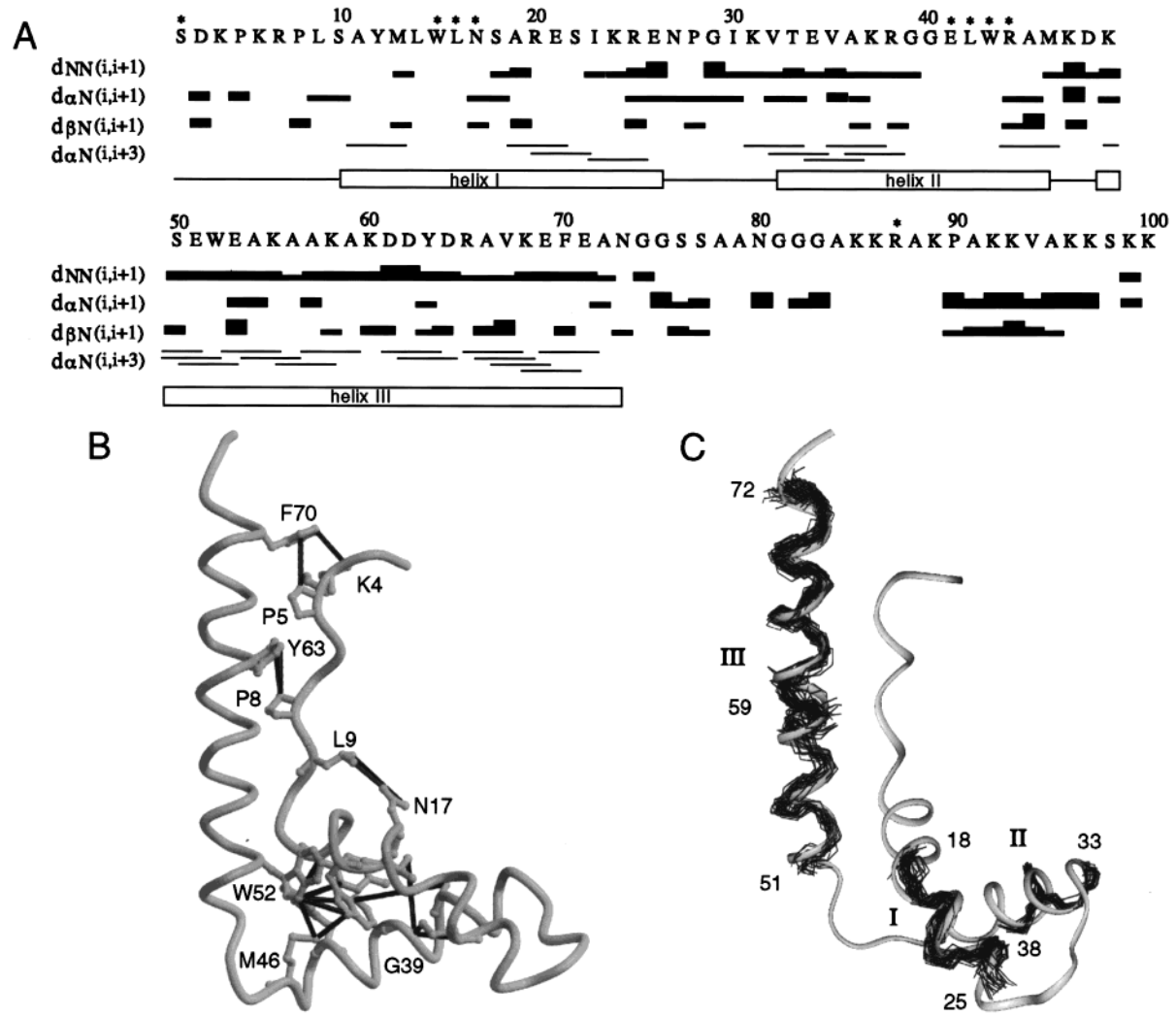

FIGURE 5: Secondary structure of the HMG-D-100 protein bound to DNA. (A) Schematic diagram of sequential NOE constraints obtained for the HMG-D-100−DNA complex. The strength of the sequential NOE interaction as strong, medium, or weak is indicated by the height of the bar. The indicated helices represent the position of α-helices in the structure of the free protein (*14*). An * denotes residues with unassigned NH groups. (B) Model of HMG-D incorporating experimental NOE results. The gray protein scaffold is based on the structure of the HMG domain of HMG-D bound to linear DNA determined crystallographically (*23*). The long-range NOEs are represented as black lines connecting residues, of which several are labeled. This figure was generated using MOLSCRIPT (*87*) and RASTER3D (*88*). (C) Model of HMG-D incorporating calculated structures of helical regions. The gray protein scaffold is based on the structure of the HMG domain of HMG-D bound to linear DNA determined crystallographically (*23*). The families of structures calculated for the helical regions of the protein are shown as black lines. This figure was generated using INSIGHT (MSI).

The structure calculations show that all of the helices are maintained in the complex with DNA, as illustrated in Table 2 and Figure 5C. Interestingly, helix III spans residues 51−72 in the structure of the complex, but it appears to bend at residue 59. Whereas the backbone RMSD for residues 49−72 is 2.59 Å, evaluation of the superimposed structures between residues 51−59 and 59−72 reveals far lower RMSD values, indicating that the two halves of helix III are well formed. Helix III is not disordered in the structures of HMG domain−DNA complexes of LEF-1, SRY, HMG1, and HMG-D (*20*, *21*, *23*), but a model of NHP6A bound to DNA (*61*) and the predicted structure of HMG-D bound to DNA (*57*) do show a bend at this position. The abundance of small hydrophobic residues combined with one discontinuity in the (*i*,*i*+3) NOEs, may permit NOE constraints in this region to be satisfied in many different conformations. Further, the local density of positively charged lysine residues, with no balancing negative charge, may also lead to helical distortions in structure calculations. Together with the CSI analysis and the NOE analysis, these results suggest that helix III is most likely to be continuous and is not significantly bent but do not rule out a kink or bend in the helix near position 59.

DNA-Binding Surface of HMG-D-100. The changes which occur in HMG-D upon binding DNA are reflected by differences in chemical shift between the free and bound forms of the protein. Both the altered electronic environment of the nucleus caused by the nearby presence of DNA and the altered conformation of the protein induced by DNA binding can influence chemical shift differences. The chemical shift difference comparison for NH, N, and Hα is shown in Figure 6A−C. The relatively large differences in Hα chemical shift near the N-terminus of the protein are explained by the unusually low shifts seen in the free protein: 2.40 ppm for Tyr12 and 3.57 ppm for Arg7. The normalized mean chemical shift difference, ($\Delta_{AVE}/\Delta_{MAX}$), per residue (Figure 6D) and the color-mapped structure of HMG-

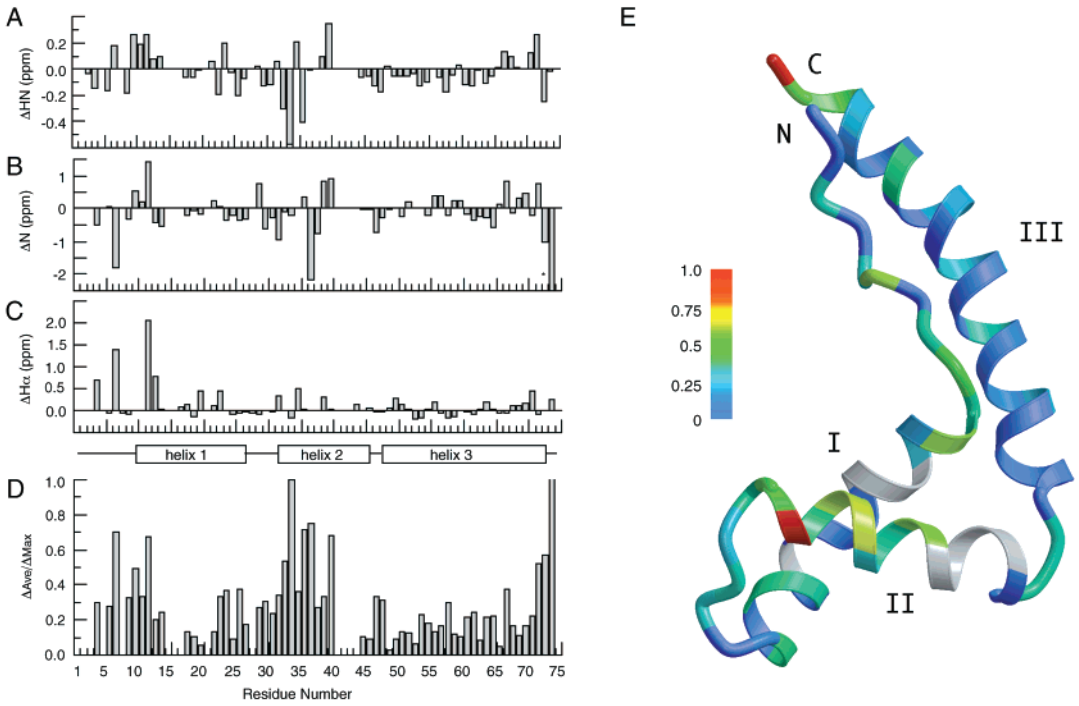

FIGURE 6: Chemical shift changes between the free and complexed forms of HMG-D. (A), (B), and (C) show the chemical shift differences in amide proton, nitrogen, and Hα, as Δ_{NH}, Δ_N, and $\Delta_{H\alpha}$, respectively. (D) shows the normalized mean shift difference ($\Delta_{AVE}/\Delta_{MAX}$) per residue, where $\Delta_{AVE} = [0.5(\Delta_{NH}^2 + \Delta_N^2/25)]^{1/2}$ (*89*). The residue with the second largest chemical shift perturbation, residue 34, rather than residue 74, was used for Δ_{MAX} in the normalization, because the complex persists for an additional 26 residues beyond the end of the free protein. Assignments are partial for residues 2, 17, and 44, are missing for residues 15, 16, and 41−43, and are tentative for residue 21. (E) Normalized mean shift differences ($\Delta_{AVE}/\Delta_{MAX}$) are color mapped onto a representative structure of free HMG-D-74. The scale ranges from blue (minimum shifts) through green and yellow to red (maximum shifts). Those residues which were unassigned or had ambiguous assignments are in white, and prolines are in blue. The helices, N-terminus, and C-terminus are labeled. This figure was generated using MOLSCRIPT (*87*) and RASTER3D (*88*).

D-74 (Figure 6E) illustrate the regions of the protein influenced most by DNA binding. The greatest changes in chemical shift occur in the loop between helix I and helix II and in the N-terminal part of helix II, along the one face of the helix that would contact the DNA. The region that displays the second greatest chemical shift differences includes residues 7−12, which is consistent with known DNA contacts and partial intercalation of Met13 (*23*).

In contrast to the first two helices, helix III shows little significant change in backbone chemical shift, indicating that this region is not interacting closely with the DNA nor has it undergone any significant conformational changes as a result of DNA binding (residues 73 and 74 were not considered because the C-terminus of the free protein was residue 74). There is no clear indication that HMG-D has interactions similar to those made by Lys53 and Arg60 in the third helix of LEF-1, and the solution data are completely consistent with the helical character and the position of helix III in the HMG-D−DNA crystal structure (*23*).

The contact regions of the HMG domain of HMG-D with DNA in solution indicated by the chemical shift analysis are consistent with the mode of DNA binding observed in the LEF-1−DNA complex in solution (*20*) and in the crystal structure of the HMG domain of HMG-D bound to linear DNA (*23*). Specifically in LEF-1, a large number of contacts cluster near the intercalating methionine 13 residue at the N-terminus of the protein at positions 6−13, 16, 17, and 20. The differences in chemical shift upon complex formation for HMG-D at residues 7, 10, and 12 indicate that similar contacts may be occurring in the non-sequence-specific protein. The second cluster of contacts between LEF-1 and DNA occurs at the amino terminus of helix II at residues 32, 33, 36, and 37, precisely where the largest chemical shift perturbations occur in this study (at residues 33, 34, 36, 37, and 40). A second intercalation event between the DNA and a residue in the loop between helices I and II of the protein, which was first predicted by Balaeff et al., occurs in the cocrystal structure of HMG-D with linear DNA (*57*) and explains the degree of change seen in these regions of the protein in this study.

The positions of the protein that were not assigned due to missing resonances were also in regions predicted to interact with the DNA. Despite the formation of a tight complex with this particular DNA fragment, as seen in binding studies, those regions of the protein that make direct contacts with the DNA appear to be exchange broadened on the NMR time scale. This result suggests that the protein is not specifically fixed to the DNA but shifts rapidly among different conformations. We have previously observed conformational averaging for HMG-D bound to linear DNA (*9*), and line broadening was reported by Allain et al. in the NMR study of NHP6A bound to DNA (*61*).

The Unstructured Basic Tail Interacts with DNA. The basic tail of HMG-D-100, residues 80−100, is important for increasing the affinity of the HMG domain by more than 200-fold, but how this increase in affinity is obtained is not known. The backbone chemical shifts of the basic tail region were easily assigned, but the NOE experiments revealed predominantly intraresidue NOEs in the basic tail (Figure 6). These NOEs were insufficient to support any regular

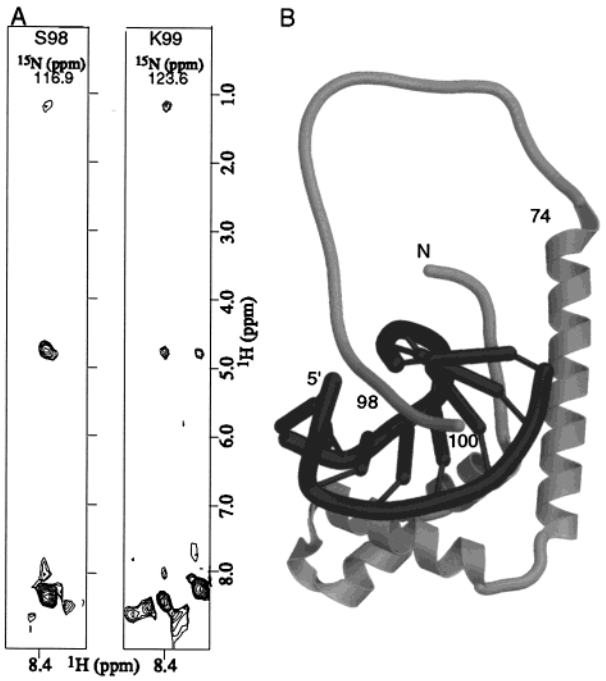

FIGURE 7: HMG-D basic tail−DNA contacts. (A) Planes of the $^{12}C^{13}C$-filtered NOESY spectrum show the backbone amide proton contacts of Ser98 and Lys99 with DNA at 1.18 and 1.13 ppm, respectively. The peaks at the bottom of the planes are diagonal peaks while those near 5 ppm represent interactions with water. (B) Model of the HMG-D−DNA complex summarizing experimental results and constraints. The gray protein scaffold is based on the structure of the HMG domain of HMG-D bound to linear DNA (in dark gray) determined crystallographically (*23*). The tail was modeled in a configuration that is consistent with the data obtained in this study. This figure was generated using MOLSCRIPT (*87*) and RASTER3D (*88*).

secondary structure or to define the orientation of the tail relative to the protein. The CSI analysis confirms that the tail adopts no regular secondary structure in this complex.

Despite the lack of secondary structure, direct contacts between two key amino acid residues in the tail and the DNA have been observed. $^{12}C^{13}C$-filtered NOESY experiments permit observation of contacts between the labeled protein and the unlabeled DNA in the natural context of the non-sequence-specific complex, without altering the system by incorporating artificial cross-links between the protein and the DNA. Figure 7A shows a portion of a $^{12}C^{13}C$-filtered NOESY spectrum, which illustrates that the amide protons of Ser98 and Lys99 make contact with DNA protons resonating at 1.18 and 1.13 ppm, respectively. Such chemical shift values are characteristic of thymine methyl groups, suggesting that the backbone of HMG-D-100's tail is contacting a thymine base, or bases, in the major groove. The sequence of DNA used in this study is 5′-CGAATCXA-TTGATTCG-3′, where X indicates the position of the special phosphoramidite. The center set of thymine bases appears to be buried behind the disulfide cross-link and is likely to be unavailable for protein−DNA contacts, but unequivocal identification of the particular thymine base or bases being contacted is not possible with the available data for this complex. The proximity of the backbone of residues 98 and 99 to the bases of the DNA indicates that the backbone inserts into the major groove and the side chains extend into solution. A model that satisfies the structural constraints of the experimental data and illustrates the basic tail interaction in the major groove is shown in Figure 7B.

DISCUSSION

These NMR studies of HMG-D-100 bound to disulfide cross-linked DNA examined the nature of non-sequence-specific DNA recognition by the HMG domain and the basic tail of HMG-D in solution. The HMG domain has structure and contact regions similar to those seen in HMG domain structures with linear DNA. NMR studies of this complex also provided a novel detailed view of the histone H1-like AK motif in a DNA complex. We show, under conditions where the tail increases DNA affinity by over 100-fold, that it is unstructured in solution. We also identify residues at the C-terminus that appear to interact with DNA in the major groove, which may be important in the regulation of DNA binding of HMG-D via phosphorylation.

HMG Domain−DNA Interactions. Previous binding studies suggested that a high-affinity structure-specific complex was formed between HMG-D-100 and DNA containing a synthetic disulfide cross-link (*62*). The studies described here indicate that residues in the interface of the complex are in intermediate exchange on the NMR time scale, suggesting that a nonspecific complex is being formed. The HMG-D−DNA cocrystal structure may explain these observations, because no specific protein−DNA hydrogen bonds are observed that could fix the protein in one conformation (*23*). Further, the disulfide cross-linked DNA appears to lack a single highly prekinked base step that could contribute to site specificity; it is also less conformationally constrained when compared with cisplatin adducts, since the structure of the free cross-linked DNA reflects an average conformation over the time period of the NMR experiment (*62*). In contrast, prekinked DNA, such as the cisplatin-modified (*22*) or a bulged DNA (*63*), has a single kinked site for which HMG domains have increased affinity and specificity since they can lodge an intercalating residue (position 32/33) into the kink (*62*). Therefore, kinked DNA might be expected to form better site-specific complexes with HMG domains than either linear DNA or the disulfide cross-linked DNA used here. The intermediate exchange observed in this structural analysis, the line-broadening observed in the NHP6A−DNA complex study (*61*), and the fast exchange observed in complexes of HMG1 with DNA (*17*) may indicate that the chromosomal HMG domain proteins do not readily form specific complexes in solution in the absence of a severe kink in the DNA.

The HMG domain of HMG-D has DNA contact regions similar to those seen in other HMG domain structures with linear DNA. The N-terminal residues surrounding the intercalating methionine and the N-terminal region of helix II likely make close contacts with DNA. This is consistent with the residues observed to interact with DNA in the NMR structure of LEF-1 with DNA and in DNA-binding studies with HMG1 (*17*), NHP6A (*61*), and HMG-D (*23*). Helix III appears to make no contacts with the DNA and does not exhibit any significant conformational change as a result of DNA binding. This is consistent with the cocrystal structure of HMG-D bound to DNA. In contrast, the mode of binding observed for HMG domains with other distorted DNA molecules is slightly different from that observed for the

linear DNA (*23*). In the structure of HMG1 bound to "distorted" cisplatin-modified DNA, the DNA recruits the protein to one side of the single drug-induced kink, and there is little interaction with the N-terminal stretch of the HMG domain. Such a structure would not account for the large number of chemical shift differences that were observed in this region of HMG-D (*22*). Comparison of the cisplatin-modified DNA to the disulfide cross-linked DNA also suggests that the interactions of HMG domain proteins with cross-linked DNA more closely resemble those with linear DNA. The chemical shift results for the solution structure of the HMG-D−disulfide cross-linked DNA complex and the structure of the substrate itself argue for a mode of DNA binding that more closely resembles that observed for the linear DNA than that for prekinked DNA.

There is a high degree of secondary structure similarity in the N-terminal residues, 1−63, between the chromosomal and transcription factor HMG domains, but the secondary structure at the C-terminus of helix III, residues 63−74, differs dramatically. Helix III terminates cleanly at proline 68 (HMG-D numbering) in the sequence-specific proteins, LEF-1 and SRY (*20*, *21*). This proline is not present in HMG-D or NHP6A where helix III extends at least six residues beyond that of LEF-1 and SRY (*14*, *20*, *21*). The sequence-specific and non-sequence-specific HMG proteins have distinctly different helix III−C-terminal tail junctions that appear to be important in directing the interaction of the tail with DNA.

The structural differences at the C-terminus of helix III lead to varying effects on both DNA bending and binding. In the structure of LEF-1 bound to DNA the kink introduced into the protein at proline 68 positions the basic tail for close contact with both minor and major grooves (*20*, *61*). The major groove interactions appear to stabilize the complex by permitting close approach of the DNA phosphate backbones that are present due to the compressed major groove. Both the N-terminal basic tail of NHP6A (*61*) and the C-terminal tail of LEF-1 (*20*), which are required for high affinity and DNA bending, bind in the major groove at the inside of the DNA bend. Whereas bending and binding are both affected by the mutation or deletion of the basic tail for NHP6A (*25*) and LEF-1 (*20*), equivalent alterations of the HMG-D basic tail only impact DNA affinity (*26*, *27*). The DNA contacts observed via serine 98 and lysine 99 with probable thymine methyl groups would also have to occur in the major groove of DNA. Interestingly, the basic tail of HMG-D is longer than that of NHP6A or LEF-1, and only the C-terminus of the basic tail is observed to occupy the major groove. The tails of the sequence-specific proteins are close enough to enter the major groove directly, but the basic tail of HMG-D is free in solution, potentially interacting with DNA loosely before entering the major groove. This difference may explain the roles of the basic tails in stabilizing DNA interactions and influencing DNA bending.

DNA Interactions of AK Motifs. The absence of regular secondary structure in the basic tail region of HMG-D is significant. The formation of segmented α-helices in the comparable tail region of histone H1 has been shown using circular dichroism (*30*). Fiber diffraction studies of model peptides and modeling studies have also suggested that the histone basic tail motifs would form helical structure in the context of DNA binding (*31*, *32*, *64*). Helical wheel diagrams of residues 83−100 of HMG-D predict that the tail forms an amphipathic helix with the basic residues on one face of the helix (not shown). However, the tail of HMG-D, which contains two AK motif heptad repeats, is not α-helical in the presence of disulfide cross-linked DNA in solution, under conditions where it enhances DNA affinity by greater than 100-fold. Further, circular dichroism studies performed in the presence and absence of linear DNA duplexes of different lengths also show no increase in helical content of the protein (data not shown). We conclude that the tail motifs interact with DNA without any regular secondary structure. In addition, it is possible that the motifs may be signals for protein modifications that modulate DNA affinity or that higher order structures may be required to induce structures that are not observed in the experiments performed here.

Differences between the interactions of the HMG-D basic tail and other HMG domain proteins have implications for DNA affinity and the attenuating effects of protein modifications. In the complexes formed between DNA and LEF-1 or NHP6A, the basic tail residues interact with the DNA in the major groove via side chain interactions, with the backbone amides positioned at such a distance that no NOEs to the bases are possible (*20*, *61*). In the HMG-D−DNA complex, the backbone amide protons of residues 98 and 99 are within 5 Å of the DNA bases, which allows the side chains in this region to extend into solution. This configuration permits modification of these residues by either phosphorylation or acetylation, providing the cell with a means of altering the affinity of HMG-D for DNA. The sequence of HMG-D in the vicinity of the basic tail that contacts the major groove is 96-KKSKK-100, a potential phosphorylation site for protein kinase C. It is known that, in vivo, HMG-D is phosphorylated (*33*, *34*) and acetylated (S. Ner, unpublished results) and that substitution of the equivalent serine with aspartic acid in a related HMG domain protein, *Chironomus* HMG1a, reduces DNA binding affinity 5-fold (*33*). Therefore, phosphorylation and/or acetylation at or near the KKSKK sequence of HMG-D could be used to attenuate the affinity of HMG-D for DNA, modulating its role in chromatin organization.

Basic tails with conserved sequences, such as the AK motif, are found in a number of chromosomal proteins that play an architectural role in DNA compaction. The functions of histone H1 proteins, for example, involve nucleosome binding and the formation of additional higher order chromatin structures. This work reveals a novel mode of binding for the AK motifs. Further, it suggests that the AK motifs of histone H1 may interact with DNA in a similar fashion as described here for HMG-D and that the mechanism by which their affinity is modulated by phosphorylation may be similar.

ACKNOWLEDGMENT

We appreciate the helpful suggestions and contributions of members of the Churchill laboratory, including Frank Murphy and Janet Klass for constructive comments on the manuscript, and Heidi Wanaski for technical assistance. We thank Alexander Balaeff for assistance with VMD and X-PLOR and Kevin Gardner for advice on expression of deuterium-labeled protein. We acknowledge the Varian Oxford Instrument Center for Excellence in NMR Laboratory at the University of Illinois at Urbana−Champaign.

REFERENCES

1. Bustin, M., and Reeves, R. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 54*, 35−100.
2. Wolffe, A. P. (1995) *Chromatin Structure and Function*, Academic Press, London.
3. van Gent, D. C., Hiom, K., Paull, T. T., and Gellert, M. (1997) *EMBO J. 16*, 1665−2670.
4. Bustin, M. (1999) *Mol. Cell. Biol. 19*, 5237−5246.
5. Bianchi, M. E., Falciola, L., Ferrari, S., and Lilley, D. (1992) *EMBO J. 11*, 1055−1063.
6. Kolodrubetz, D., and Burgum, A. (1990) *J. Biol. Chem. 265*, 3234−3239.
7. Wagner, C. R., Hamana, K., and Elgin, S. C. R. (1992) *Mol. Cell. Biol. 12*, 1915−1923.
8. Ner, S. S., and Travers, A. A. (1994) *EMBO J. 13*, 1817−1822.
9. Churchill, M. E. A., Jones, D. N. M., Glaser, T., Hefner, H., Searles, M. A., and Travers, A. A. (1995) *EMBO J. 14*, 1264−1275.
10. Giese, K., Amsterdam, A., and Grosschedl, R. (1991) *Genes Dev. 5*, 2567−2578.
11. Harley, V. R., Jackson, D. I., Hextall, P. J., Hawkins, J. R., Berkovitz, G. D., Sockanathan, S., Lovell-Badge, R., and Goodfellow, P. N. (1992) *Science 255*, 453−456.
12. Teo, S. H., Grasser, K. D., and Thomas, J. O. (1995) *Eur. J. Biochem. 230*, 943−950.
13. Grosschedl, R., Giese, K., and Pagel, J. (1994) *Trends Genet. 10*, 94−100.
14. Jones, D. N., Searles, M. A., Shaw, G. L., Churchill, M. E., Ner, S. S., Keeler, J., Travers, A. A., and Neuhaus, D. (1994) *Structure 2*, 609−627.
15. Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A., Laue, E. D., and Thomas, J. O. (1993) *EMBO J. 12*, 1311−1319.
16. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., and Norman, D. G. (1993) *Nucleic Acids Res. 21*, 3427−3436.
17. Hardman, C. H., Broadhurst, R. W., Raine, A. R. C., Grasser, K. D., Thomas, J. O., and Laue, E. D. (1995) *Biochemistry 34*, 16596−16607.
18. van Houte, L. P. A., Chuprina, V. P., van der Wetering, M., Boelens, R., Kaptein, R., and Clevers, H. (1995) *J. Biol. Chem. 270*, 30516−30524.
19. King, C.-Y., and Weiss, M. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11990−11994.
20. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature 376*, 791−795.
21. Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) *Cell 81*, 705−714.
22. Ohndorf, U.-M., Rould, M. A., He, Q., Pabo, C. O., and Lippard, S. J. (1999) *Nature 399*, 708−712.
23. Murphy, F. V., Sweet, R. M., and Churchill, M. E. A. (1999) *EMBO J. 18*, 6610−6618.
24. Lnenicek-Allen, M., Read, C. M., and Crane-Robinson, C. (1996) *Nucleic Acids Res. 24*, 1047−1051.
25. Yen, Y. M., Wong, B., and Johnson, R. C. (1998) *J. Biol. Chem. 273*, 4424−4435.
26. Payet, D., and Travers, A. (1997) *J. Mol. Biol. 266*, 66−75.
27. Churchill, M. E. A., Changela, A., Dow, L. K., and Krieg, A. J. (1999) *Methods Enzymol. 304*, 99−133.
28. Wells, D., and McBride, C. (1989) *Nucleic Acids Res. 17*, 311−347.
29. Allan, J., Hartman, P. G., Crane-Robinson, C., and Aviles, F. X. (1980) *Nature 288*, 675−679.
30. Clark, D. J., Hill, C. S., Martin, S. R., and Thomas, J. O. (1988) *EMBO J. 7*, 69−75.
31. Azorin, F., Vives, J., Campos, J. L., Jordan, A., Lloveras, J., Puigjaner, L., Subirana, J. A., Mayer, R., and Brack, A. (1985) *J. Mol. Biol. 185*, 371−387.
32. Subirana, J. A. (1990) *Biopolymers 29*, 1351−1357.
33. Wisniewski, J. R., and Schulze, E. (1994) *J. Biol. Chem. 269*, 10713−10719.
34. Wisniewski, J. R., Szewczuk, Z., Petry, I., Schwanbeck, R., and Renner, U. (1999) *J. Biol. Chem. 274*, 20116-2012.
35. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.
36. Dow, L., Changela, A., Hefner, H., and Churchill, M. (1997) *FEBS Lett. 414*, 514−520.
37. Ferentz, A. E., Keating, T. A., and Verdine, G. L. (1993) *J. Am. Chem. Soc. 115*, 9006−9014.
38. Plateau, P., and Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.
39. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
40. Geen, H., and Freeman, R. (1991) *J. Magn. Reson. 93*, 93−141.
41. Emsley, L., and Bodenhausen, G. (1992) *J. Magn. Reson. 97*, 135−148.
42. Kupce, E., Boyd, J., and Campbell, I. D. (1995) *J. Magn. Reson., Ser. B 106*, 300−303.
43. Baum, J., Tycko, R., and Pines, A. (1985) *Phys. Rev. A 32*, 3435−3448.
44. Silver, M. S., Joseph, R. I., and Hoult, D. I. (1984) *J. Magn. Reson. 59*, 347−351.
45. Kupce, E., and Freeman, R. (1997) *J. Magn. Reson. 127*, 36−48.
46. Shaka, A. J., Lee, C. J., and Pines, A. (1988) *J. Magn. Reson. 77*, 274−293.
47. Shaka, A. J., Keeler, J., Frenkiel, T., and Freeman, R. (1983) *J. Magn. Reson. 52*, 335−338.
48. Shaka, A. J., Barker, P. B., and Freeman, R. (1985) *J. Magn. Reson. 64*, 547−552.
49. McCoy, M. A., and Mueller, L. (1993) *J. Magn. Reson. A101*, 122−130.
50. Levitt, M. H., and Freeman, R. (1981) *J. Magn. Reson. 43*, 502−507.
51. Cavanagh, J., Fairbrother, W. J., Palmer, A. G. I., and Skelton, N. J. (1996) *Protein NMR spectroscopy principles and practice*, Academic Press, San Diego.
52. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J. Biomol. NMR 6*, 135−140.
53. Brunger, A. T. (1992) *X-PLOR 3.1, A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
54. Nilges, M. (1993) *Proteins: Struct., Funct., Genet. 17*, 297−309.
55. Fletcher, C. M., Jones, D. N. M., Diamond, R., and Neuhaus, D. (1996) *J. Biomol. NMR 8*, 292−310.
56. Diamond, R. (1992) *Protein Sci. 1*, 1279−1287.
57. Balaeff, A., Churchill, M. E. A., and Schulten, K. (1998) *Proteins: Struct., Funct., Genet. 30*, 113−135.
58. Humphrey, W., Dalke, A., and Schulten, K. (1996) *J. Mol. Graphics 14*, 33−38.
59. Venters, R. A., Farmer, B. T. I., Fierke, C. A., and Spicer, L. D. (1996) *J. Mol. Biol. 264*, 1101−1116.
60. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., Carrillo, M. O. M., and Norman, D. G. (1995) in *Nucleic Acids and Molecular Biology* (Ekstein, F., and Lilley, D. M. J., Eds.) pp 222−250, Springer-Verlag, Berlin.
61. Allain, F. H.-T., Yen, Y.-M., Masse, J. E., Schultze, P., Dieckmann, T., Johnson, R. C., and Feigon, J. (1999) *EMBO J. 18*, 2563−2579.
62. Wolfe, S. A., Ferentz, A. E., Grantcharova, V., Churchill, M. E. A., and Verdine, G. L. (1995) *Chem. Biol. 2*, 213−221.
63. Payet, D., Hillisch, A., Lowe, N., Diekmann, S., and Travers, A. (1999) *J. Mol. Biol. 294*, 79−91.
64. Churchill, M. E. A., and Travers, A. A. (1991) *Trends Biochem. Sci. 16*, 92−97.
65. Huth, J. R., Bewley, C. A., Nissen, M. S., Evans, J. N. S., Reeves, R., Gronenborn, A. M., and Clore, G. M. (1997) *Nat. Struct. Biol. 4*, 657−665.
66. Bodenhausen, G., and Ruben, D. (1980) *Chem. Phys. Lett. 69*, 185−187.
67. Bax, A., and Ikura, M. (1991) *J. Biomol. NMR 1*, 99−104.
68. Grzesiek, S., and Bax, A. (1992) *J. Magn. Reson. 96*, 432−440.
69. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.

70. Clubb, R. T., Thanabal, V., and Wagner, G. (1992) *J. Magn. Reson. 97*, 213−217.
71. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 11655−11666.
72. Wittekind, M., and Meuller, L. (1993) *J. Magn. Reson., Ser. B 101*, 201−205.
73. Grzesiek, S., and Bax, A. (1993) *J. Biomol. NMR 3*, 185−204.
74. Powers, R., Gronenborn, A. M., Clore, G. M., and Bax, A. (1991) *J. Magn. Reson. 94*, 209−213.
75. Vuister, G. W., Wang, A. C., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 5334−5335.
76. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
77. Kay, L. E., Ikura, M., and Bax, A. (1991) *J. Magn. Reson. 91*, 84−92.
78. Fesik, S. W., and Zuiderweg, E. R. (1990) *Q. Rev. Biophys. 23*, 97−131.
79. Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1992) *J. Biomol. NMR 2*, 655−659.
80. Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1992) *FEBS Lett. 314*, 413−418.
81. Grzesiek, S., Anglister, J., Ren, H., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 4369−4370.
82. Montelione, G. T., Lyons, B. A., Emerson, S. D., and Tashiro, M. (1992) *J. Am. Chem. Soc. 114*, 10974−10975.
83. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
84. Zhang, O., Forman-Kay, J. D., Shortle, D., and Kay, L. E. (1997) *J. Biomol. NMR 9*, 181−200.
85. Ikura, M., and Bax, A. (1992) *J. Am. Chem. Soc. 114*, 2433−2440.
86. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
87. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
88. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
89. Grzesiek, S., Bax, A., Clore, G. M., Gronenborn, A. M., Hu, J.-S., Kaufman, J., Palmer, I., Stahl, S. J., and Wingfield, P. T. (1996) *Nat. Struct. Biol. 3*, 340−345.

BI000723V