Determining the DNA Bending Angle Induced by Non-Specific High Mobility Group-1 (HMG-1) Proteins: A Novel Method†

Lingjuan Tang, Jin Li, Darin S. Katz, and Jin-an Feng*

Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Received August 4, 1999; Revised Manuscript Received November 24, 1999

ABSTRACT: To study the DNA bending induced by non-sequence-specific HMG-1 domain proteins, we have engineered a fusion protein linking the yeast NHP6A with a sequence-specific DNA binding domain, the DNA binding domain of the Hin recombinase, Hin-DBD. A series of biochemical experiments were carried out to characterize the DNA binding property of this fusion protein. Our data showed that the fusion protein not only specifically recognizes a DNA fragment containing the Hin-DBD binding site, but also binds DNA with a higher affinity in comparison with either domain alone. Both domains of the fusion protein are bound to the DNA in juxtaposition. Permutation assays showed that the fusion protein induced a DNA bending at the site of NHP6A binding by an estimated value of 63°. We believe that this experimental design provides an effective vehicle to determine the DNA bending induced by nonspecific HMG-1 proteins.

One of the more exciting findings of the non-histone chromatin proteins is the discovery of a large group of specialized proteins, otherwise unrelated, containing a segment of sequence that is homologous to the high mobility group 1 (HMG-1)[1] domain (*1*, *2*). These proteins function in different reactions of cellular processes, including DNA replication, transcription, and recombination. Lacking a distinct activation or enzymatic domain, the HMG-1-like proteins are mainly structural proteins. In vitro experiments have shown that this group of proteins plays an important functional role in engineering active nucleoprotein complexes (*1*, *3*, *4*). UBF, a nucleolar RNA polymerase I transcription factor, is one of the best-characterized mammalian HMG-1-like proteins (*5*). An electron spectroscopy imaging study has shown that a dimer of UBF transcription factor can wrap more than 180 base pairs (bp) of the promoter DNA, bringing two distant regulatory elements into proximity for optimal interaction and promoting transcription (*6*). Yeast HMG-1-like proteins, NHP6A/B, have been shown to facilitate promoter-specific transcriptional activation in vivo by mediating the assembly of preinitiation complex involving proteins bound to the distantly spaced TATA-box and UAS sites (*7*). Other studies have shown that HMG-1/2 proteins can also promote V−(D)−J recombination catalyzed by RAG 1 and RAG 2 in vitro (*8*) and p53-mediated transactivation in vivo by stimulating the assembly of higher order p53 nucleoprotein structures (*9*).

The functional roles of HMG-1 box proteins to architect unusual DNA structures are closely related to their DNA binding activities. These proteins have an especially high binding affinity to a variety of non-B-DNA structures, such as four-way DNA junction (4WJ), cruciform DNA, and prebent DNA structures (*10*−*12*). They also induce severe distortions in the regular DNA structure, such as sharp kinks, DNA looping, and DNA supercoiling (*6*, *13*−*16*). DNA ligase-mediated DNA circularization experiments showed that yeast NHP6A/B can induce a drastic DNA deformation which facilitates DNA fragments of 66 bp to form small circles (*17*). While most HMG-1 domain proteins bind DNA nonspecifically, a small group of these proteins do recognize specific DNA sequences (*11*, *18*). Biochemical and structural studies of these proteins have provided insight into how HMG-1 proteins interact with DNA.

The three-dimensional structures of DNA complexes of two sequence-specific HMG-1 domain proteins, SRY, the testes-determining factor, and LEF-1, a lymphoid enhancer-determining protein, showed that the boomerang-shaped HMG-1 domain induces large DNA bends by penetrating deep into the minor groove of the DNA (*20*, *21*). The binding of the HMG-1 domain causes the minor groove to be widened and underwound, consequently exposing a surface that accommodates the concave surface of the HMG-1 domain (*19*−*21*). Interestingly, although similar contacts are observed in both structures, the two proteins induce drastically different DNA bending angles. The distortions of the DNA helix by SRY and LEF-1, leading to an overall bending of 80° and 120°, respectively, are mainly a result of partial intercalations of amino acid side chains between base pairs of the DNA in the minor groove (*20*, *21*). In light of the very nature of the architectural function of HMG-1-like proteins, these varying degrees of DNA distortions may potentially have significant biological implications.

Molecular characterization of DNA binding by non-sequence-specific HMG-1 domain proteins, on the other

† This work was supported in part by National Institutes of Health Grant GM54630 (J.F.), American Cancer Society Grant PRG9926301GMC (J.F.), and an appropriation from the Commonwealth of Pennsylvania. D.S.K. is supported by a postdoctoral fellowship from NIH.

* To whom correspondence should be addressed. Phone: (215) 728-5306. Fax: (215) 728-3105. E-mail: feng@guanyin.fccc.edu.

[1] Abbreviations: HMG, high-mobility group; bp, base pair; EMSA, electrophoretic mobility shift assays; IPTG, isopropyl β-D-thiogalactopyranoside; LEF-1, lymphoid enhancer binding factor 1; NHP6A, nonhistone protein 6A.; SRY, sex-determining region Y protein; tris, tris(hydroxymethyl)aminomethane.

10.1021/bi991819w CCC: $19.00
Published on Web 02/25/2000

hand, is limited. Current molecular biology and biophysical techniques are unable to produce results that detail how each individual nonspecific HMG-1 domain interacts with DNA. The main difficulty of studying nonspecific DNA binding proteins at the molecular level is the inevitable generation of heterogeneous species in solution. Chow and Lippard (*22*) have developed a method that localizes the nonspecific HMG-1-like protein onto a DNA probe with a preferred binding site created by cisplatin cross-linking. By cloning the prebent DNA probe into a permutation plasmid, they were able to estimate bending angles induced by a number of HMG-1-like proteins (*22*). However, the biological relevance of the artificially engineered binding site is still unclear.

We describe here a novel method to characterize the DNA binding property of a non-sequence-specific HMG-1 protein from *Saccharomyces cerevisiae*, NHP6A. This approach involves constructing a fusion protein in which the NHP6A is linked with a sequence-specific DNA binding domain, the DNA binding domain of the Hin recombinase, Hin-DBD. The rationale of this design is for the Hin-DBD to "bring" the NHP6A protein to a specific DNA site in juxtaposition to the Hin-DBD binding site, and result in a one-to-one homogeneous protein/DNA complex. A series of experiments were carried out to characterize the DNA binding property of the Hin-DBD/NHP6A fusion protein. Electrophoretic mobility shift assays (EMSA, also known as gel retardation assay or gel shift assay) were carried out to test whether the fusion protein binds DNA specifically. The kinetics of DNA binding affinity were analyzed by the method of surface plasmon resonance biosensor (BIAcore, Pharmacia). DNase I footprinting assays were performed to evaluate whether both domains of the fusion protein bind the DNA in juxtaposition. Finally, DNA permutation assays were carried out to evaluate the bending angle induced by the fusion protein. These experiments demonstrate that this fusion protein not only binds DNA sequence specifically, but also functions as an efficient vehicle that allows us to determine the bending angle induced by nonspecific HMG proteins.

EXPERIMENTAL PROCEDURE

Design and Construction of the Fusion Proteins. In an attempt to design a fusion protein for the binding study, we considered two factors that are critical to the success of this experiment: (1) the specific DNA binding domain should have a stable fold and form a specific complex with DNA as a monomer; (2) the linker sequence bridging the two domains of the fusion protein should be long and flexible enough, preferably adopting an extended conformation, so that potential "interference binding" between the domains of the fusion protein is minimized.

The DNA binding domain of the Hin recombinase appears to meet our criteria sufficiently (*23*, *24*). The Hin recombinase mediates a site-specific DNA recombination that controls the alternate expression of two flagella surface antigens in *Salmonella* (*25*). By specifically recognizing the recombination site, which contains a sequence characteristic of an inverted repeat, the Hin cleaves the DNA at the center of its binding site in a staggered fashion, and relegates the DNA after strand exchange (*25*). Although the Hin recombinase functions as a dimer, its DNA binding domain (Hin-DBD, residues 139−190) can recognize half of the inverted

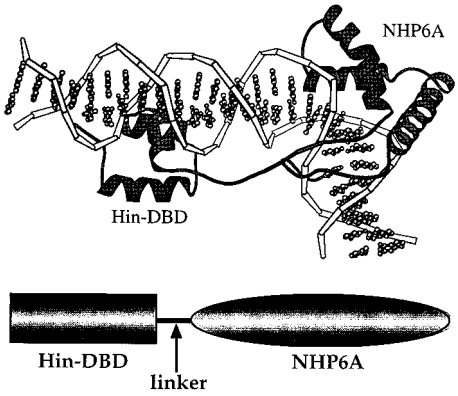


FIGURE 1: (a, top) A computer model of the Hin-DBD/NHP6A fusion protein bound to DNA. (b, bottom) Schematic of the fusion protein design. All components are labeled.

repeat (half-site) as a monomer with moderately strong affinity ($K_d \approx 10^{-7}$ M) (*26*). The crystal structure of the Hin-DBD bound to a 14-bp DNA oligomer containing the half-site shows that the Hin-DBD folds into a three-α-helical structure with extended polypeptide chains on both the N- and C-terminals of the domain (*24*) (Figure 1a). The Hin-DBD recognizes its binding site by inserting a helix−turn−helix (HTH) DNA binding motif into the major groove of the DNA, establishing specific protein−DNA contacts. Both the N- and C-terminal peptides are also involved in interacting with the DNA (*24*). The N-terminal arm is inserted into the minor groove and makes specific hydrogen bonds with one of the A−T base pairs in the minor groove. The C-terminal tail, while not significantly affecting the DNA binding affinity of the Hin-DBD (*26*), is found to be positioned in the minor groove in an extended conformation (*23*). This feature is particularly favorable to our design of the fusion protein, as it could function as a "natural" linker bridging Hin-DBD and NHP6A. Adding a short linker of 3−4 residues that fuses NHP6A to the C-terminal of Hin-DBD could result in a domain separation by as long as 3.5 Å × 11 = 38.5 Å, which should provide ample flexibility for "noninterference" binding to the DNA by either domain.

Production and Purification of Hin-DBD/NHP6A Fusion Protein. DNA fragments encoding Hin-DBD (residues 139−190) and NHP6A were generated by PCR. The PCR products of both genes were fused together with linker sequence that contains staggered ends matching the 3′ end of the Hin-DBD gene on one end to the 5′ end of the NHP6A gene on the other. Three pieces of DNA fragments were then linked together with DNA ligase, and subcloned into an expression vector, pET11a (Novagen). The plasmid was then transformed into BL21 (DE3) bacterial strain for protein overexpression. The sequence of fusion gene was confirmed by DNA sequencing using ABI 377 automatic DNA sequencer. Figure 1b shows a schematic of the domain structure of the fusion protein that contains the Hin-DBD on the N-terminal end, a linker of three amino acids, Glu-Phe-Arg, and the NHP6A on the C-terminal. The fusion protein expression was induced with 1.0 mM IPTG for 2 h at 37 °C. The cells were harvested and stored at −70 °C.

For protein purification, the cells were thawed and resuspended in 7.0 mL of lysis buffer containing 20 mM Tris−HCl (pH 7.5), 5% glycerol, 1.0 mM DTT, 1.0 mM benzamidine, and 0.1 mM EDTA (pH 8.0) for every gram

of wet cells. The cells were then lysed by French Press. The bacterial lysate was centrifuged at 14 000 rpm for 45 min at 4.0 °C. The supernatant was used for subsequent purification. NaCl was added to adjust its concentration to 1.0 M. DNA was removed by the addition of a sufficient amount of polyethylenimine (PEI; Sigma) to generate maximum precipitate in the high-salt extract followed by centrifugation at 12 000 rpm for 30 min. Remaining PEI in the supernatant was removed by batch chromatography using 0.2 volume of phosphocellulose (Whatman P11, pH 8.0) slurry. The lysate was then dialyzed into a low-salt buffer (LSB) containing 20 mM Tris−HCl (pH 7.5), 75 mM NaCl, 5% glycerol, 1 mM DTT, 1.0 mM benzaminide, and 0.1 mM EDTA (pH 8.0). The dialyzed protein solution was then passed through a cation-exchange (SP-Sepharose Fast Flow, Pharmacia Biotech) column which was equilibrated with the same buffer. The fusion protein was eluted with a linear gradient from 75 mM to 1.5 M NaCl in LSB. Fractions containing the fusion protein were collected and concentrated to a concentration of 2.0 mg/mL. The concentrated protein sample was passed through a gel filtration column (Superdex 75 HR, Pharmacia) using FPLC. The fusion protein was eluted in the buffer of 20 mM Tris−HCl, 500 mM NaCl, 5% glycerol, 1 mM DTT, 1 mM benzaminide, and 0.1 mM EDTA (pH 8.0). Fractions containing the fusion protein were concentrated with Centricon (Amicon Inc.) to a final concentration of 1.0 mg/mL in the same buffer. The purity of the fusion protein used for subsequent DNA binding assays was at least 85% on the basis of Coomassie blue stained acrylamide gel (data not shown). The molecular weight of the fusion protein was confirmed by mass spectroscopy, which showed a dominant peak of molecular weight 20 826 (unpublished results), consistent with the predicted value based on amino acid composition of the fusion protein (molecular weight 20 847).

Gel Shift Assays. A DNA oligomer of 36 bp (36-mer) of the sequence (5′-gcgTTTTTGATAAGAATgggctgaatcgagctggtc-3′), where capital letters indicate the Hin-DBD binding site, was used for electrophoretic mobility shift assay. Binding reactions were performed in 1X binding buffer containing 30 mM Tris-HCl (pH 7.5), 125 mM NaCl, 0.004% NP-40 2% glycerol, 0.1 mM EDTA (pH 8.0). A 15 nM concentration of ^{32}P-labeled DNA was incubated with increasing concentrations of proteins at room temperature for 50 min. Identical volumes of samples were loaded onto 8% polyacrylamide gels (29:1 acrylamide/bisacrylamide), which were prerun for 30 min in 1X TBE buffer. Electrophoresis was performed for 2.5 h at 14 mA. Gels were dried and exposed to X-ray film. Competition assays were performed with the addition of competing oligomers. Details of these experiments are described in the Results. The results of gel shift assays are shown in Figures 2−4.

DNA Binding Assay with a BIAcore System. The oligonucleotide and the fusion protein used in the binding assays with a BIAcore1000 were in the buffer containing 20 mM HEPES−KOH (pH 7.5), 80 mM KC1, 5.0 mM $MgCl_2$, 1.0 mM DTT, and 0.005% surfactant P-20. A 3′-biotinylated 36-mer which contains the Hin-DBD binding site was immobilized on a streptavidin-conjugated biosensor chip (BIAcore) by injecting 325 μL of the 0.5 M oligonucleotide solution at 10 μL/min followed by 200 μL of 5.0 μM free biotin. The biosensor chip carrying the upper strand of

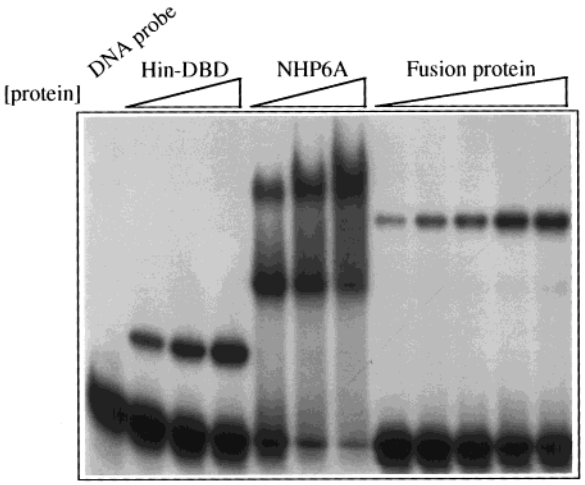


FIGURE 2: Gel shifts on a 36-bp linear DNA fragment by Hin-DBD, NHP6A, and the fusion protein. ^{32}P-labeled 36-bp DNA fragment (15 nM) was incubated with buffer alone or with increasing amounts of Hin-DBD (1−3 ng), or NHP6A (2−8 ng) and the fusion protein (2−32 ng) in increments of 2-fold.

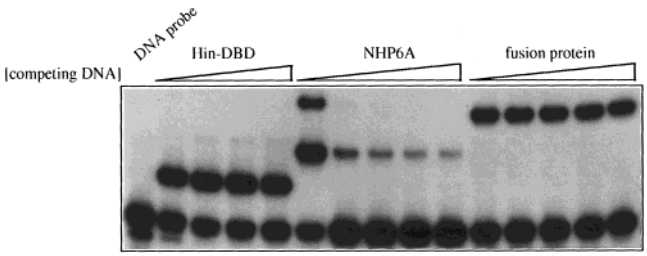


FIGURE 3: Gel shifts on the 36-bp DNA fragment by Hin-DBD, NHP6A, and the fusion protein in a concentration-dependent challenging assay. ^{32}P-labeled 36-bp (15 nM) was preincubated with proteins (Hin-DBD, 3 ng; NHP6A, 4 ng; fusion protein, 16 ng) before excess amounts of competing DNA oligomers were added: 0-, 100-, 200-, 300-fold excess to the Hin-DBD/36-bp mixture; 0-, 50-, 100-, 200-, 300-fold excess to the NHP6A/36-bp mixture; and 0-, 50-, 100-, 200-, 300-fold excess to the fusion protein/36-bp mixture.

oligonucleotide was used repeatedly through assay cycles consisting of the following four steps: (a) injection of 15 μL of 2.0 μM lower-strand oligonucleotide at 2.0 μL/min; (b) injection of 250 μL of various concentrations of proteins (either NHP6A or the fusion protein) at 30 μL/min followed by 100 μL of the buffer at the same flow rate; (c) injection of 30 μL of 1.0 M KCl containing buffer at 30 μL/min; (d) injection of 30 μL of 30 mM NaOH at 30 μL/min. Dissociation constants (K_d) were calculated from the ratio of the dissociation rate (k_{dissoc}) and association rate (k_{assoc}), which were derived from the measured relative response unit (RU) on the basis of a one-to-one interaction model by using the software BIAevaluation 2.0 (BIAcore). The results are shown in Figure 5 and Table 1.

Dnase I Footprinting. A 36-bp double-stranded oligonucleotide containing the Hin-DBD binding site was cloned into pBluescript II SK (Stratagene) plasmid at the Sma I site. Following digestion with Xho I, the DNA was labeled with [α-^{32}P]dTTP using the Klenow fragment of DNA polymerase I. After further digestion with BssH II, a 175-bp DNA fragment containing the Hin-DBD binding site was isolated and purified using conventional nondenaturing polyacrylamide gel electrophoresis. Binding reactions were performed in 1X binding buffer [20 mM Tris−HCl (pH 7.6), 100 mM

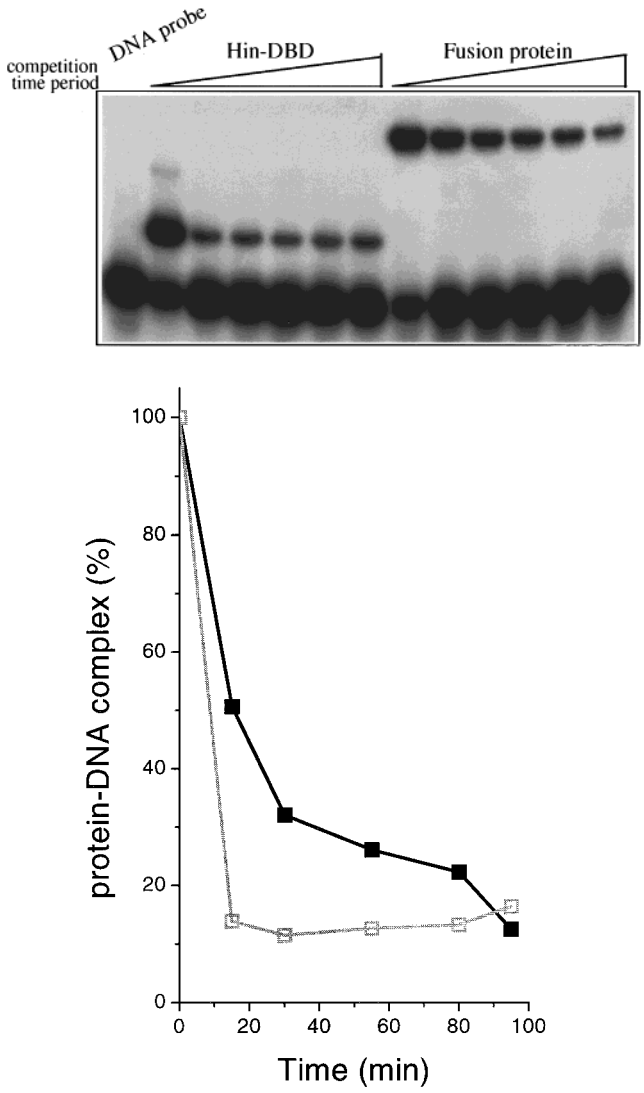


FIGURE 4: (a, top) Gel shifts on the 36-bp DNA fragment by Hin-DBD and the fusion protein in a time course competition experiment. ^{32}P-labeled 36-bp (15 nM) was preincubated with the proteins (Hin-DBD, 3 ng; fusion protein, 16 ng). A 150-fold molar excess of competing 36-mer oligomer was added to the mixture and allowed to compete for the time periods of 15, 30, 55, 80, and 90 min. (b, bottom) Analysis of the dissociation rate. The curve connecting open squares represents data of the Hin-DBD/DNA complex; the curve connecting filled squares represents data of the fusion protein/DNA complex.

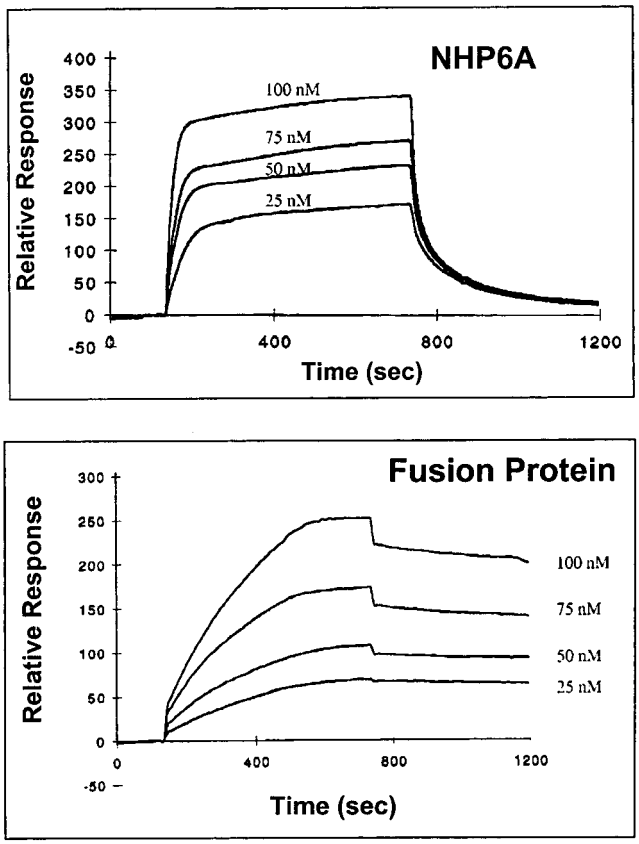


FIGURE 5: Profiles of BIAcore analysis of the binding affinity of NHP6A and the fusion protein. The relative response is plotted against the time course of the experiments. Concentrations of proteins used in each experiment are indicated. (a, top) Binding of the NHP6A to immobilized 36-bp on the biosensor chip. (b, bottom) Binding of the fusion protein to immobilized 36-bp.

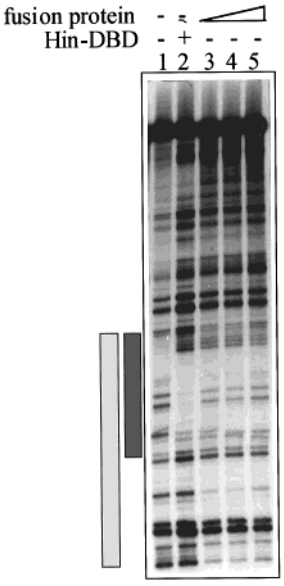
Table 1: Dissociation Rate Constant for NHP6A and the Fusion Protein

	fusion protein		NHP6A	
concn (nM)	k_{assoc} ($\times 10^4$ M^{-1} s^{-1})	k_{dissoc} ($\times 10^{-4}$ s^{-1})	k_{assoc} ($\times 10^5$ M^{-1} s^{-1})	k_{dissoc} ($\times 10^{-2}$ s^{-1})
25	11.6 ± 0.1	0.94 ± 0.1	2.1 ± 0.1	1.0 ± 0.1
50	6.7 ± 0.1	1.2 ± 0.1	2.9 ± 0.1	1.1 ± 0.1
75	5.5 ± 0.1	1.9 ± 0.1	2.3 ± 0.1	1.2 ± 0.1
100	3.5 ± 0.1	1.9 ± 0.1	2.4 ± 0.1	1.3 ± 0.1
	$K_d = 2.2 \times 10^{-9}$ M		$K_d = 4.8 \times 10^{-8}$ M	

NaCl, 5.0 mM $MgCl_2$, 1.0 mM DTT, 10 μg/mL polycytidylic acid] in 45 μL volumes for 20 min. Following incubation of protein and DNA, 5.0 μL of 2.5 μg/mL DNase I was added and the reaction was incubated for an additional 30 s. A 6.0 μL volume of cold stop solution [666 μM Tris−HCl (pH 9.5), 66 μM CDTA, 3.3% SDS] was added. The digested DNA was subsequently extracted and isolated using one phenol/chloroform extraction. Digested DNA was redissolved in deionized formamide containing xylene cyanol FF and bromophenol blue as dye markers. Samples were loaded onto 6.0% (w/v) polyacrylamide/7.5 M urea gels. Gels ran at 1500 V for 3 h, after which they were dried and exposed to X-ray film (Figure 6a).

Circular Permutation Assays. Developed by Wu and Crothers (*27*), the circular permutation assay is a commonly used electrophoretic assay to evaluate DNA bending. The principle of this assay is based on the experimental observation that the electrophoretic mobility of a DNA fragment in a polyacrylamide gel is correlated with the bending locus on the DNA fragment. A DNA fragment with a sharp bend at the center migrates slower than a DNA fragment with a bend near its termini. We cloned a 36-bp double oligonucleotide containing the Hin-DBD binding site into a circular permutation vector pCY4, in which a number of specific endonuclease restriction sites were engineered (Figure 7). The Hin-DBD binding site was inserted between SST I and Bgl II sites on the pCY4 vector. This plasmid was digested with five restriction enzymes, *Eco*RI, *Hin*dIII, *Eco*RV, *Nhe*I, and *Bam*HI, generating five DNA fragments for binding assay (Figure 7). These DNA fragments were then labeled with [α-^{32}P]dTTP after purification.

Gel mobility shift assays were performed by incubating fusion protein with 25 000 cpm end-labeled DNA probe in 10 μL solution containing Tris−HCl (pH 7.5), 125 mM

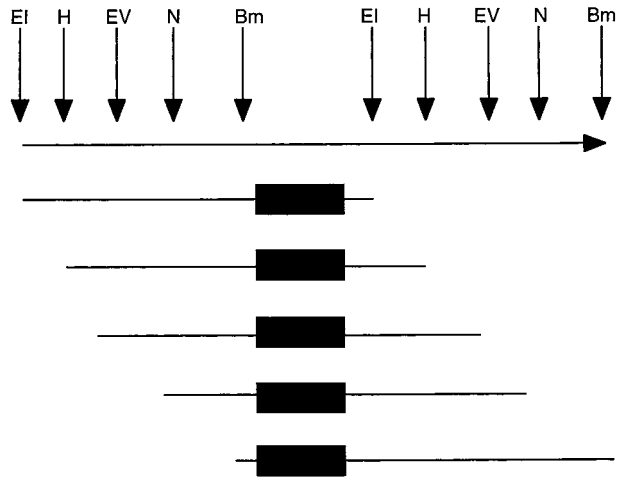


Summary of Protection Results

	GCGTTTTTGATAAGAATGGGCTAGAATCG (plus strand)
Hin-DBD	+ **** *
fusion protein	**** * * *

\+ band is darker when protein is bound
* band is lighter or disappeared when protein is bound

FIGURE 6: (a, top) Comparative footprinting of Hin-DBD and the fusion protein across the inserted Hin-DBD binding site. All lanes are labeled. Lane 1 is the unprotected DNA probe. The location of the Hin-DBD binding site is indicated by a dark gray rectangular bar; the DNA sequence covered by the fusion protein is marked by a light gray rectangular bar. Protein concentration used in each reaction: Hin-DBD, 10 μM; fusion protein, 0.5, 1.0, and 5.0 μM. (b, bottom) Summary of the protection results.



Key: EI=EcoRI; H=HindIII; EV=EcoRV; N=NheI; Bm=BamHI

FIGURE 7: Plasmid, pCY4, containing the Hin-DBD binding site (filled box), was cleaved at the restriction sites indicated in the map. Abbreviations used for the restriction enzymes are indicated at the bottom of the figure. The DNA fragments obtained in this way all contain circular permutations of the same sequence 412-bp.

NaCl, 2.0% glycerol, 0.04% NP-40, 0.1 mM EDTA (pH 8.0), and 15 ng Poly dI/dC for 50 min at 25 °C. A 2.5 μL volume of 7.5% Ficoll plus 0.1% bromphenol blue in the above buffer was then added. After 5 min the reaction mixtures were loaded onto 6% polyacrylamide gel (40:1 acrylamide/bisacrylamide in 0.5X TBE) that was electrophoresing at 7.5 V/cm for about 1.5 h. Electrophoresis proceeded at 15 V/cm for 3 h at 4.0 °C, and the locations and relative amount of the bound and unbound forms were determined by autoradiography or with a PhosphorImager (Figure 8a). The gel mobility assay shown in Figure 8b was performed under the same buffer conditions, except more poly dI/dC (5 μg) was added to the reaction mixture and the gel was electrophoresed at 15 V/cm for 6.0 h at 4.0 °C.

The bending angles were estimated as described by Thompson and Landy (*28*). Briefly, Hin-NHP6a complexes formed on the restriction fragment from pCY4 derivatives digested with *Eco*RI (Hin binding site at one end) and *Eco*RV (Hin binding site at the center of the fragment) were electrophoresed as above (Figure 8a). End-labeled fragments containing from two to seven in-phase A-tracts located at the center and one end of the restriction fragments (from pjT170-2 through pjT170-7 provided by R. C. Johnson, UCLA) were electrophoresed in parallel (Figure 9a). A plot of the relative mobility (μ_M/μ_E) of the A-tract standards, the ratio of the migration of the band corresponding to the site at the middle of the fragment (μ_M) divided by the migration of the fragment when the binding site is at the end (μ_E), was generated (Figure 9b). The number of "A-tract equivalents" was then determined for the relative mobility of the fusion protein from the plot, and the angles were calculated on the basis of the value of 18° per A-tract (28).

RESULTS

The Hin-DBD/NHP6A fusion protein binds DNA sequence specifically. To address the DNA binding specificity of the fusion protein, we designed and synthesized a 36-bp oligonucleotide containing the sequence of 5′-(gcgTTTTTG-ATAAGAATgggctgaatcgagctggtc)-3′, where capital letters indicate the sequence of the half Hin recombination site (Hin-DBD binding site). The 36-mer was then labeled with [γ-^{32}P]-ATP. Figure 2 shows the result of an electrophoretic mobility shift assay on the Hin-DBD, the NHP6A, and the fusion protein binding to the labeled 36-mer. As can be seen, the electrophoretic migration patterns of DNA complexes of Hin-DBD and NHP6A are markedly different. The Hin-DBD forms a homogeneous complex with a single retarded electrophoretic migration band, while NHP6A forms multiple migration bands as the protein concentration is increased, typical for a nonspecific DNA binding protein. The gel shift pattern of the fusion protein, on the other hand, is similar to that of the Hin-DBD, i.e., forming a single stable complex with the target site.

The relative DNA binding specificity of the fusion protein in comparison with both Hin-DBD and NHP6A was also addressed by competition challenge assays. In a concentration-dependent competition assay, the labeled 36-bp oligonucleotide was preincubated with a predetermined amount of proteins before being challenged with an excess amount of unlabeled 36-bp DNA oligomer of same sequence (36-mer). As shown in Figure 3, the prebound complexes formed with Hin-DBD, NHP6A, and the fusion protein were each challenged with 50-, 100-, 200-, and 300-fold molar excesses of competing oligomer. The preformed Hin-DBD/36-mer complex was not affected by the increasing concentrations of the competing DNA. The NHP6A/36-mer complexes were

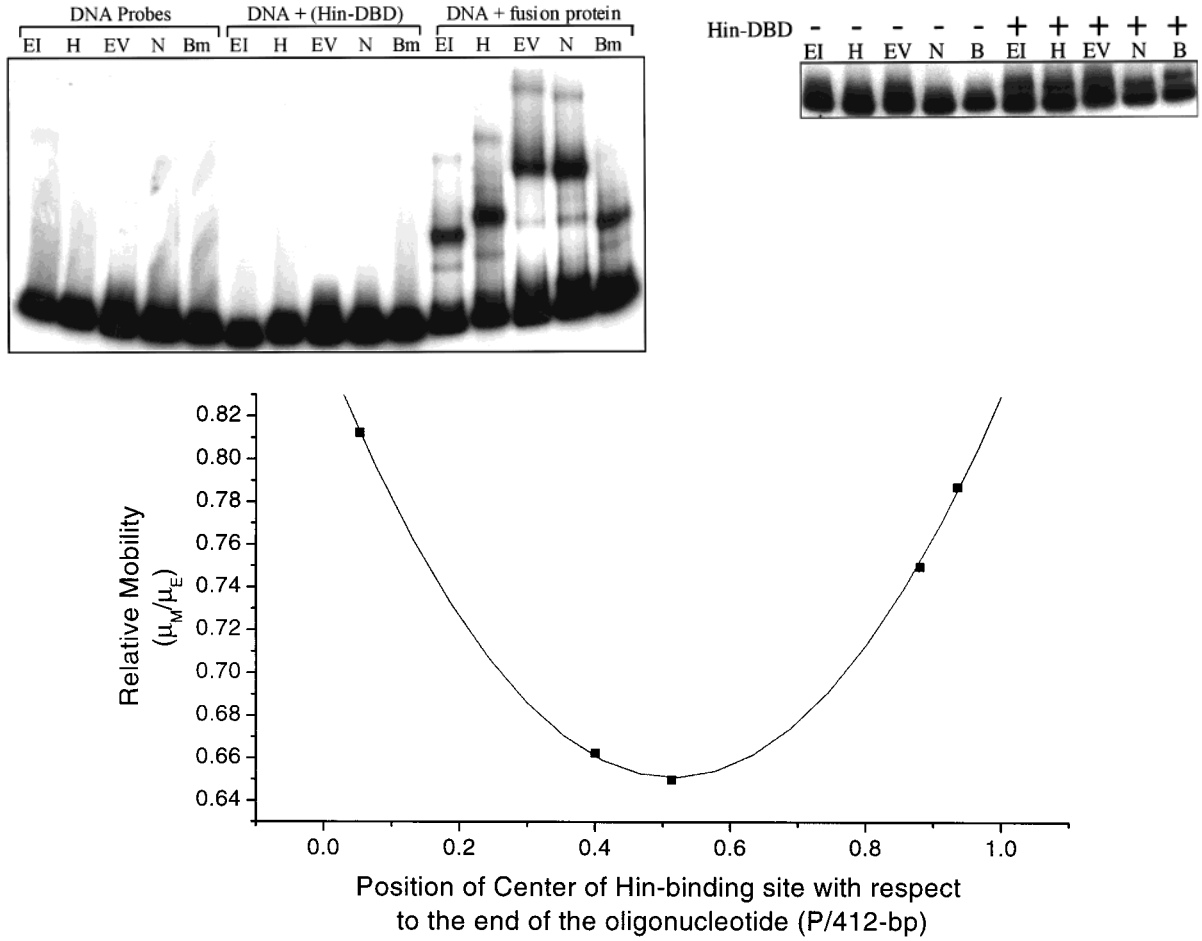


FIGURE 8: Circular permutation analysis of the DNA bending induced by the binding of the fusion protein. (a, top) Gel shifts using each of the circular permutated DNA fragments, generated with different restriction enzyme cuts, complexed with the fusion protein. Each lane is labeled according to the restriction enzyme used to generate the DNA fragment. Abbreviations are the same as those used in Figure 7. The DNA probes used in the reaction are shown in the left panel (five lanes), the five lanes in the middle panel are Hin-DBD/DNA complexes, and the five lanes in the right panel are fusion protein/DNA complexes. (b, middle) Permutation DNA fragments complexed with the Hin-DBD. The five lanes on the left show DNA probes as control; Hin-DBD/DNA complexes are shown on the right. (c, bottom) Analysis of the induced DNA bends. The relative mobilities of the protein/DNA complexes are plotted against the location of the Hin-DBD binding site on the fragment. The curve displayed is a least-squares fit of the data points.

not quite as stable. In the presence of a 50-fold molar excess of the competitor, the NHP6A/36-mer complex band started to weaken significantly. When the competitor concentration reached 100-fold, the NHP6A/36-mer complex almost completely dissociated. In contrast, the fusion protein has a DNA binding selectivity similar to that of the Hin-DBD. The prebound fusion protein remained complexed with the 36-mer even in the presence of a 300-fold molar excess of the competing oligonucleotides.

The stability of the fusion protein/DNA complex in comparison with that of the Hin-DBD was addressed by a time course competition gel shift assay where the preformed complexes were incubated with a 150-fold molar excess of unlabeled 36-mer for different periods of time. A marked difference in dissociation rates between the two complexes was observed (Figure 4). In the 15 min time period, at least 50% of the fusion protein stayed bound to the labeled 36-mer, while almost 90% of the Hin-DBD was already dissociated from the labeled 36-mer. The fusion protein remained bound until after 80 min of competition. It is apparent that the dissociation rate for the fusion protein/DNA is slower than that of the Hin-DBD/DNA complex.

The Fusion Protein Binds Its Target Site with High Affinity. To determine specific binding affinity of the fusion protein with its target site, we carried out a kinetic binding study with a surface plasmon resonance biosensor (BIAcore, Pharmacia), which analyzes the association and the dissociation rates of complexes formed between molecules in the solution and molecules prebound to a sensor chip by measuring the reflective index change in units of response unit (RU). In this experiment, the 36-mer oligonucleotide carrying the Hin-DBD binding site was immobilized onto the surface of the carboxymethylated dextran sensor chip. Four different concentrations of NHP6A and the fusion protein (25, 50, 75, and 100 nM) were applied to the reaction. Parts a and b of Figure 5 show profiles of binding experiments for NHP6A and the fusion protein, respectively. The dissociation constants for both proteins were calculated on the basis of the values of association and dissociation rates. The nature of both the association and the dissociation of the DNA with NHP6A appeared to be quite different from those with the fusion protein. The NHP6A associated with DNA quickly and reached a steady state at both high and low concentrations of the protein; the average k_{assoc} for the four experiments of different protein concentrations is $(2.4 \pm 0.1) \times 10^5$ M^{-1} s^{-1}. In contrast, the association of the fusion protein is slower. At low protein concentrations, the response increased gradually, and did not reach equilibrium

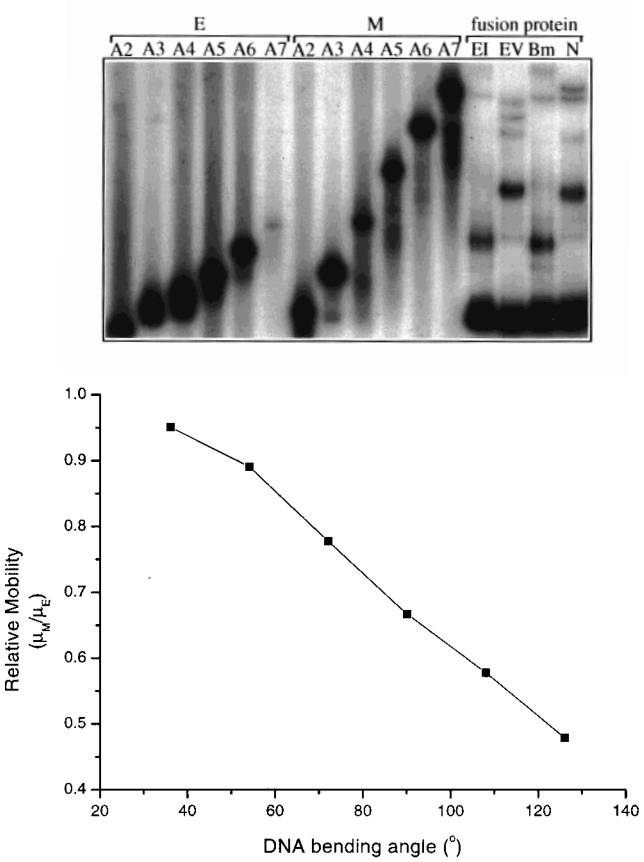


FIGURE 9: (a, top) Gel shifts of A-tract sequences located either on the end (E) or in the middle (M) of the DNA fragments. Labels A2−A7 indicate the number of in-phase A-tract sequences in the DNA fragment; i.e., A2 indicates two in-phase A-tracts. The four lanes in the right panel are gel shifts of the fusion protein/412-bp complex with a Hin-DBD binding site on the ends (EI and Bm) and in the middle (EV and N) (see Figure 7 for details). (b, bottom) A standard curve of DNA bending vs the ratio of relative mobility (see the text).

within the time period of the experiment. Only when the concentration of the fusion protein increased above 75 nM did the association reach a steady state. The association pattern observed here apparently is typical for sequence-specific DNA binding proteins (*29*, *30*). The average k_{assoc} for four measurements at different fusion protein concentrations is $(6.8 \pm 0.1) \times 10^4$ M^{-1} s^{-1}. Showing characteristics of nonspecific DNA binding proteins, the NHP6A dissociated quickly as soon as the protein concentration dropped off. The average k_{dissoc} is $(1.2 \pm 0.1) \times 10^{-2}$ s^{-1}. On the other hand, the dissociation of the fusion protein is significantly slower. Most of the fusion protein stayed bound even after the protein concentration in the solution dropped off to zero. The average k_{dissoc} for the fusion protein is estimated at $(1.48 \pm 0.1) \times 10^{-4}$ s^{-1}.

On the basis of these experiments, the DNA binding constant for NHP6A is estimated at 48 ± 1.0 nM, which is comparable to what had been reported previously, ~100 nM (*17*). The binding constant for the fusion protein is 2.2 ± 0.1 nM. Therefore, in comparison with the DNA binding affinity of Hin-DBD (100 nM) (*26*) and that of the NHP6A (50−100 nM), the fusion protein not only binds specifically to the DNA oligomer containing the Hin-DBD binding site, but also has higher DNA binding affinity.

Both Domains of the Fusion Protein Interact with the DNA. The experimental data described above demonstrated that the fusion protein has the DNA binding characteristics of a sequence-specific DNA binding protein. To address the question of whether the NHP6A domain of the fusion protein interacts with the DNA in juxtaposition to the designed Hin-DBD binding site, the DNAase I cleavage protection assay was employed. Figure 6a shows the DNase I protection patterns produced by both the Hin-DBD and the fusion protein.

The Hin-DBD, which specifically binds to the Hin-DBD binding site, protected a region of approximately 13 bp (Figure 6a, lane 2) as compared to the control (lane 1), in which no protein was added. This result is consistent with previous DNA footprinting assays on Hin-DBD and the intact Hin recombinase (*31*, *32*). The crystal structure of the Hin-DBD/DNA complex shows that the HTH DNA binding motif interacts with the central GATA of the Hin-DBD binding site, and that the N-terminal arm and the C-terminal tail of the Hin-DBD strengthen the DNA binding by interacting with minor grooves of the sequences flanking both ends of the central GATA sequence (*24*). The DNase I digestion of the fusion protein/DNA complex showed a protection or enhancement of cleavage that covered two regions of approximately 25 bp (lanes 3−5). The extended coverage beyond the 13-bp Hin-DBD binding site suggests direct contacts between NHP6A and the DNA. A summary of the extent of DNA protection by both the Hin-DBD and the fusion protein is shown in Figure 6b. It is clear that both domains of the fusion protein bind the DNA in juxtaposition, in which the Hin-DBD binds the Hin-DBD binding site and the NHP6A binds a segment of DNA 3′ to it. NHP6A alone, which does not have any specificity for DNA sequences, produces no distinct protection pattern (Katz and Feng, unpublished results).

The Fusion Protein Induces DNA Bending. To demonstrate whether the NHP6A domain of the fusion protein is actively engaged in DNA binding, i.e., inducing DNA bending, we carried out circular permutation assays. Figure 8a shows a representative gel of this experiment. The left panel (five lanes) of the gel shows the five DNA probes, the middle panel (five lanes) shows the Hin-DBD complex, and the right panel (five lanes) of the gel shows the fusion protein complex with each DNA fragment. The pattern of the DNA migration in this retardation gel clearly shows DNA bending induced by the fusion protein, while Hin-DBD causes no bend in DNA. The gel shifts of the Hin-DBD/DNA complex were only observable after electrophoresing for an extended period of time (Figure 8b). On the other hand, the rate of DNA migration for the fusion protein/DNA complex varies with the location of the Hin-DBD binding site on the DNA fragment. The minor bands in lanes of the fusion protein/DNA complex are a result of nonspecific protein−DNA interaction, as the intensity of these bands weakens significantly with the addition of increasing amounts of poly dI/dC in the reaction mixture (data not shown). The locus of the DNA bend by the fusion protein is determined by plotting the relative mobilities of the complexes as a function of the distance from the 5′ end of the Hin-DBD binding site to the restriction enzyme site. The data were then fit by nonlinear least-squares analysis (KaleidaGraph version 3.0). The maximum revealed that the bend center occurred at the site between base pairs 205−210, which corresponds to a site next to the Hin-DBD binding site when it is positioned at

the center of the permutation fragment (Figure 8c). It is evident that the observed DNA bending is a result of NHP6A binding.

The NHP6A Protein Bends the DNA by 63°. The extent of DNA bending induced by the fusion protein can now be estimated according to an empirical method described by Thompson and Landy (*28*). The fusion protein/DNA complexes were electrophoresed alongside a set of DNA standards containing different numbers of A-tracts (A2−A7, where the number represents the number of in-phase A-tract sequences inserted into the DNA fragment; i.e., A2 represent two A-tracts) located either on the end or in the middle of the DNA fragments (Figure 9a) (*28*). A standard curve is generated by plotting the ratio of relative mobility of DNA with an A-tract bend in the middle and at the end (μ_M/μ_E) against the predicted bending angle obtained assuming each A-tract causes an 18° bend (Figure 9b). The ratio of the relative mobility (μ_M/μ_E) for the DNA bound with the fusion protein is 0.82, which corresponds to a value of 63 ± 2° in the standard curve (Figure 9b). The DNA bending angle can also be estimated on the basis of Figure 8a,c. The plotted points in Figure 8c were fitted to a second-order polynomial equation, to give $y = 0.852x^2 - 0.781x + 0.760$ ($r^2 = 0.9966$) (*11*). The induced DNA bend was determined from the first- and second-order parameters to yield an averaged value of ∼62°.

DISCUSSION

We have demonstrated here that the fusion protein can serve as an effective vehicle to determine the DNA bending induced by non-sequence-specific DNA binding proteins. The fusion protein of Hin-DBD/NHP6A exhibited DNA binding characteristics of both domains: the Hin-DBD specifically targets its binding site, and the NHP6A binds the neighboring DNA sequence, inducing a large DNA bend. The fusion protein showed a marked increase in DNA binding affinity by a factor of 25−50-fold in comparison with either domain alone. This increased binding affinity of the fusion protein is consistent with results reported for another engineered DNA binding fusion protein. Kim et al. reported a significant increase in DNA binding affinity when a TATA-box-binding protein and a zinc-finger peptide were fused together (*33*). Given the experimental data presented in this paper, it is unlikely that the high DNA binding affinity of the Hin-DBD/NHP6A is simply a result of increased electrostatic interactions.

The engineered fusion protein effectively converted the NHP6A into part of a sequence-specific DNA binding protein. Gel shift assays clearly showed that the Hin-DBD/NHP6A fusion protein specifically recognizes DNA containing the Hin-DBD binding site. The fusion protein stays bound to the Hin-DBD binding site even in the presence of a 300-fold molar excess of competing DNA (Figures 3 and 4). Furthermore, the fusion protein showed a slower dissociation rate than that of the Hin-DBD, suggesting that the NHP6A domain of the fusion protein is actively engaged in DNA binding.

The positioning of the NHP6A on the DNA is of critical importance to this fusion method. Figure 1a shows a computer model of the fusion protein bound to DNA derived from the structures of HinDBD/DNA and LEF-1/DNA complexes (*20*, *23*). The model shows that Hin-DBD binds to its binding site while bringing NHP6A onto the DNA in the minor groove. The binding of the NHP6A subsequently forces the DNA to bend toward the major groove (Figure 1a). The length of the linker sequence connecting two domains of the fusion protein plays an essential role in modulating the binding of the NHP6A. An optimal linker could limit the degree of freedom of the NHP6A without affecting its DNA binding. Structural knowledge of the HMG-1-like proteins provides an advantage in designing the linker sequence. One can build a structural model and estimate the special separation necessary for the two domains not interfering with each other when bound to the DNA. Two possible scenarios of potential interference binding could be considered: direct physical contacts between the two domains, and interference through the DNA, i.e., the binding of the one domain would distort the structure of the DNA, and subsequently affect the DNA binding of the other domain. The first scenario can be easily avoided since Hin-DBD and NHP6A bind to different grooves of the DNA (*20*, *21*, *23*). The second scenario is more complicated. There are no experimental data here that directly address this issue. The crystal structure of the Hin-DBD/DNA complex showed little DNA structural deviation from the regular B-DNA (*23*). The potential main source of interference is then the binding of NHP6A to the DNA, which could weaken the affinity of the Hin-DBD for its target site. This scenario, however, can also be avoided. Experimentally determined structures of the protein/DNA complexes, including that of the LEF-1/DNA complex, showed that large DNA bends induced by proteins were usually a result of discreet local kinks, and more significantly that the sequence flanking these kinks still adopted regular B-DNA structure (*20*, *21*, *34*−*36*). In order words, the local DNA structural distortions do not usually propagate to the neighboring sequences. On the basis of this analysis, we think it is possible to minimize potential interference binding by an insertion of a linker sequence that could allow the NHP6A to bind a DNA site 4−5 base pairs away from the Hin-DBD binding site. The experimental results described in this paper demonstrate the feasibility of this design strategy.

While circular permutation assay is one of the most widely used techniques to evaluate DNA bending, the result of DNA bending angle analysis has been shown to vary depending on the experimental conditions, such as the presence of divalent metal ions (*37*), and particularly the specific types of protein−DNA interactions (*38*, *39*). Nevertheless, the DNA bending angles determined for hSRY and mLEF-1 by the circular permutation assays were consistent with the subsequent structural determinations by NMR (*11*, *20*, *21*, *40*). In the absence of other experimental data, it is difficult to evaluate the accuracy of the bending angle determined here for the NHP6A. Preliminary NMR data suggest that the protein−DNA interactions of the NHP6A/DNA complex are likely to be similar to those of the mLEF-1/DNA complex (*41*). It is therefore reasonable to assume that the circular permutation assays could yield dependable measurements on DNA bending induced by NHP6A. Obviously, the fusion method described here can easily adopt other techniques for DNA bending angle determination, including ligase-mediated cyclization assay (*42*) and biophysical techniques such as fluorescence resonance energy transfer (FRET) (*43*). Further

studies are necessary to evaluate the accuracy of the bending angle determined here.

The DNA sequence "available" for the NHP6A binding in our construct is a GC-rich region. It has been speculated that the HMG-1 domains prefer binding to the AT-rich region of the DNA, since the minor groove of these sequences can accommodate protein insertion, establishing contacts with the floor of the minor groove. A GC-rich region, which has an amine group on the C2 position of guanine, could not facilitate such interactions. The measured 63° in this experiment could reflect a "partial" induction of the DNA bending, where the NHP6A could not completely penetrate the minor groove and establish contacts such as those observed in the structures of SRY/DNA and LEF-1/DNA complexes (*20*, *21*). Therefore, it is not surprising that the value determined in this set of experiments for the NHP6A is in the lower range of the values determined for other HMG-1 proteins (*20*, *21*). A study of the binding of this fusion protein with an AT-rich sequence in juxtaposition to the Hin-DBD binding site is already underway. It will be interesting to learn whether this target site would yield a larger DNA bend.

The significance of this fusion method is that it facilitates an opportunity to characterize DNA deformation by architectural proteins at the molecular level. It is apparent that different HMG-1 domain proteins induce different degrees of DNA bending. The differences in this DNA bending may play an important role in cellular reactions involving HMG-1 proteins. We are only beginning to learn biological functions of HMG-1 proteins. The fusion method discussed here can be a useful tool to further our understanding of the architectural function of the HMG-1 proteins.

ACKNOWLEDGMENT

We thank Reid C. Johnson and Michael Haykinson for numerous insightful discussions that led to the initiation of this work, Chiquito J. Crasto for model building and helpful comments on the manuscript, and Ms. Kathy Buchheit for her help in the preparation of the manuscript.

REFERENCES

1. Grosschedl, R., Giese, K., and Pagel, J. (1994) *Trends Genet. 10*, 94−100.
2. Bustin, M., and Reeves, R. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 54*, 35−100.
3. Crothers, D. (1993) *Curr. Biol. 3*, 675−676.
4. Sheridan, P. L., Sheline, C. T., Cannon, K., Voz, M. L., Pazin, M. J., Kadonaga, J. T., and Jones, K. A. (1995) *Genes Dev. 9*, 2090−104.
5. Jantzen, H. M., Admon, A., Bell, S. P., and Tjian, R. (1990) *Nature 344*, 830−6.
6. Bazett-Jones, D. P., Leblanc, B., Herfort, M., and Moss, T. (1994) *Science 264*, 1134−7.
7. Paull, T. T., Carey, M., and Johnson, R. C. (1996) *Genes Dev. 10*, 2769−81.
8. van Gent, D. C., Hiom, K., Paull, T. T., and Gellert, M. (1997) *EMBO J. 16*, 2665−70.
9. Jayaraman, L., Moorthy, N. C., Murthy, K. G., Manley, J. L., Bustin, M., and Prives, C. (1998) *Genes Dev. 12*, 462−72.
10. Bianchi, M. E., Beltrame, M., and Paonessa, G. (1989) *Science 243*, 1056−9.
11. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1992) *EMBO J. 11*, 4497−506.
12. Pöhler, J. R. G., Norman, D. G., Bramham, J., Bianchi, M. E., and Lilley, D. M. (1998) *EMBO J. 17*, 817−26.
13. Bianchi, M. E., and Lilley, D. M. (1995) *Nature 375*, 532.
14. Segall, A. M., Goodman, S. D., and Nash, H. A. (1994) *EMBO J. 13*, 4536−48.
15. Stros, M., Stokrova, J., and Thomas, J. O. (1994) *Nucleic Acids Res. 22*, 1044−51.
16. Stros, M., Reich, J., and Kolibalova, A. (1994) *FEBS Lett. 344*, 201−6.
17. Paull, T. T., and Johnson, R. C. (1995) *J. Biol. Chem. 270*, 8744−54.
18. Read, C. M., Cary, P. D., Preston, N. S., Lnenicek-Allen, M., and Crane-Robinson, C. (1994) *EMBO J. 13*, 5639−46.
19. Bewley, C. A., Gronenborn, A. M., and Clore, G. M. (1998) *Annu. Rev. Biophys. Biomol. Struct. 27*, 105−31.
20. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature 376*, 791−5.
21. Werner, M. H., Huth, J. R., Gronenborn, A. M., and Clore, G. M. (1995) *Cell 81*, 705−14.
22. Chow, C. S., Whitehead, J. P., and Lippard, S. J. (1994) *Biochemistry 33*, 15124−30.
23. Feng, J. A., Johnson, R. C., and Dickerson, R. E. (1994) *Science 263*, 348−55.
24. Feng, J.-A., Dickerson, R., and Johnson, R. (1994) *Curr. Opin. Struct. Biol. 4*, 60−66.
25. Johnson, R. C. (1991) *Curr. Opin. Genet. Dev. 1*, 404−11.
26. Sluka, J. P., Horvath, S. J., Glasgow, A. C., Simon, M. I., and Dervan, P. B. (1990) *Biochemistry 29*, 6551−61.
27. Wu, H. M., and Crothers, D. M. (1984) *Nature 308*, 509−13.
28. Thompson, J. F., and Landy, A. (1988) *Nucleic Acids Res. 16*, 9687−705.
29. Terenzi, H., Petropoulos, I., Ellouze, C., Takahashi, M., Zakin, M. M. (1995) *FEBS Lett. 369*, 277−82.
30. Seimiya, M., Kurosawa, Y. (1996) *FEBS Lett. 398*, 279−84.
31. Bruist, M. F., Horvath, S. J., Hood, L. E., Steitz, T. A., and Simon, M. I. (1987) *Science 235*, 777−80.
32. Glasgow, A. C., Bruist, M. F., and Simon, M. I. (1989) *J. Biol. Chem. 264*, 10072−82.
33. Kim, J. S., and Pabo, C. O. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 2812−7.
34. Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) *Nature 365*, 520−7.
35. Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) *Nature 365*, 512−20.
36. Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991) *Science 253*, 1001−7.
37. Brukner, I., Susic, S., Dlakic, M., Savic, A., Pongor, S. (1994) *J. Mol. Biol. 236*, 26−32.
38. Hagerman, P. L. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 9993−6.
39. Kerppola, T. K., and Curran, T. (1993) *Mol. Cell. Biol. 13*, 5479−89.
40. Giese, K., Cox, J., and Grosschedl, R. (1992) *Cell 69*, 185−195.
41. Allain, F. H., Yen, Y. M., Masse, J. E., Schultze, P., Dieckmann, T., Johnson, R. C., Feigon, J. (1999) *EMBO J. 18*, 2563−79.
42. Kahn, J. D., Crothers, D. M. (1998) *J. Mol. Biol. 276*, 287−309
43. Heyduk, E., Heyduk, T., Claus, P., Wisniewski, J. R. (1997) *J. Biol. Chem. 272*, 19763−19770.

BI991819W