DNA Bending Induced by High Mobility Group Proteins Studied by Fluorescence Resonance Energy Transfer[†]

Mike Lorenz,[‡] Alexander Hillisch,[‡] Dominique Payet,[§,||] Memmo Buttinelli,[§,⊥] Andrew Travers,*[,§] and Stephan Diekmann[‡]

Institute for Molecular Biotechnology, Beutenbergstrasse 11, D-07745 Jena, Germany, and Laboratory of Molecular Biology, Medical Research Council, Hills Road, Cambridge CB2 2QH, U.K.

Received February 26, 1999; Revised Manuscript Received June 7, 1999

ABSTRACT: The HMG domains of the chromosomal high mobility group proteins homologous to the vertebrate HMG1 and HMG2 proteins preferentially recognize distorted DNA structures. DNA binding also induces a substantial bend. Using fluorescence resonance energy transfer (FRET), we have determined the changes in the end-to-end distance consequent on the binding of selected insect counterparts of HMG1 to two DNA fragments, one of 18 bp containing a single dA_2 bulge and a second of 27 bp with two dA_2 bulges. The observed changes are consistent with overall bend angles for the complex of the single HMG domain with one bulge and of two domains with two bulges of ∼90−100° and ∼180−200°, respectively. The former value contrasts with an inferred value of 150° reported by Heyduk et al. (*1*) for the bend induced by a single domain. We also observe that the induced bend angle is unaffected by the presence of the C-terminal acidic region. The DNA bend of ∼95° observed in the HMG domain complexes is similar in magnitude to that induced by the TATA-binding protein (80°), each monomeric unit of the integration host factor (80°), and the LEF-1 HMG domain (107°). We suggest this value may represent a steric limitation on the extent of DNA bending induced by a single DNA-binding motif.

High mobility group (HMG)[1] proteins of the HMG1/2 class are a family of eukaryotic DNA binding and bending proteins. The archetypes of these proteins are the mammalian HMG1 and HMG2, which contain two related HMG domains, box A and box B. Homologues containing only a single HMG domain have been found in plant, yeast, and insects. The HMG box is a conserved domain of approximately 75 amino acid residues (for reviews see refs *2* and *3*). This domain has been shown by NMR spectroscopy to fold in solution into a characteristic L-shaped structure containing three α-helices with an angle of approximately 80° between the helices 2 and 3 (*4*−*6*).

The HMG domain is found in two classes of proteins: (i) the chromosomal HMG1/2 homologues which exhibit little sequence selectivity and (ii) sequence-specific transcription factors including sex-determining region Y (SRY) protein and lymphoid enhancer-binding factor (LEF-1). Both types of proteins bend DNA to a substantial extent as shown by their ability to facilitate the circularization of short DNA fragments and by gel electrophoresis (*7*−*9*). In addition, they have been proposed to act as architectural DNA-binding proteins which determine the configuration of DNA in multiprotein complexes (*10*).

The angles of the DNA bend induced by binding of SRY and LEF-1 have been determined both by circular permutation assay (*10*, *11*) and by NMR spectroscopy (*12*, *13*). SRY binds DNA in the minor groove and bends the helix by 70−80° toward the major groove; LEF-1 bends the DNA by 107−127°, and it has been reported that a basic region C-terminal to the HMG domain increases the bend angle (*14*). In contrast, the extent of bending induced by the HMG domains of the abundant chromosomal HMG proteins is uncertain, and its determination is complicated by the lack of sequence-specific binding. However, these proteins bind selectively to distorted DNA structures such as four-way junctions (*15*), bulged DNA (*16*), and *cis*-platinated DNA (*17*). The sequences flanking the HMG domain at the C-terminus modulate the affinity of the protein for DNA (*16*). The basic region of cHMG1a and HMG-D increases and the acidic tail decreases the affinity for linear DNA (*16*, *18*). However, the latter increases the selectivity for distorted DNA (*16*).

In this work we have analyzed the influence of the highly homologous proteins HMG-D and HMG-Z from *Drosophila*, cHMG1a from *Chironomus*, and NHP6A from *Saccharomyces cerevesiae* (all of which contain an HMG domain structurally homologous to the B box of HMG1) on the DNA conformation by fluorescence resonance energy transfer

[†] This work was supported by a fellowship to D.P. from the European Commission (Contract ERBFMBICT961777) and by exchange grants from the British Council and the DAAD.

* Corresponding author.

[‡] Institute for Molecular Biotechnology.

[§] Laboratory of Molecular Biology, Medical Research Council.

[||] Current address: Centre d'Immunologie de Marseille Luminy, Parc Scientifiique et Technologique de Luminy, Case 906, 13288 Marseille Cedex 9, France.

[⊥] Current address: Dipartimento di Genetica e Biologia Molecolare, Università "La Sapienza", P.le A.Moro 5, 00185 Roma, Italy.

[1] Abbreviations: FRET, fluorescence resonance energy transfer; HMG, high mobility group; LEF-1, lymphoid enhancer-binding factor 1; SRY, sex-determining region Y protein; NMR, nuclear magnetic resonance; TMRh, 5-carboxytetramethylrhodamine; FAM, 6-carboxyfluorescein; TBE, 90 mM Tris−borate and 2 mM ethylenediaminetetraacetic acid, pH 8.3.

10.1021/bi990459+ CCC: $18.00
Published on Web 08/21/1999

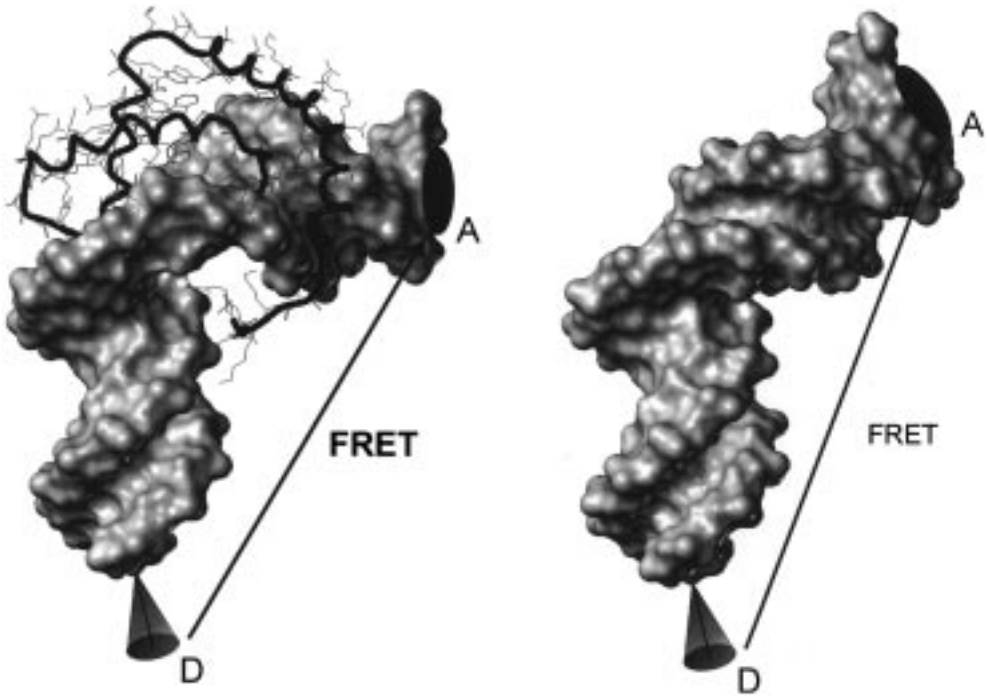

FIGURE 1: Principle of measurement of distances by FRET. The figure shows the distances between the donor and acceptor fluorophores in the HMG-D−bulged DNA complex and in the free bulged DNA. The distance between the fluorophores in the complex (thick arrow) is shorter than that in the free DNA, resulting in a higher FRET efficiency in the complex. The distributions of the 200 energetically most favorable conformers for 6-carboxyfluorescein (cone) and TMRh at the DNA helix ends (ellipse) are shown. D = donor fluorescein; A = acceptor tetramethylrhodamine.

(FRET). FRET is an excellent method to measure distances in the range from 10 to 100 Å between dye molecules covalently bound to biomolecules as DNA or RNA (Figure 1). A particularly appropriate application is the detection of conformational changes, e.g., the influence of bulges on DNA and RNA structure (*19*) DNA bending by TATA-binding protein (TBP) (*20*) and by PAP1 (*21*). We have exploited the property of preferred binding of the HMG domain to distorted DNA structures and have chosen DNA sequences containing dA_2 bulges as selective binding sites (*16*). We have used two DNA fragments, one of 18 bp containing a single dA_2 bulge and a second of 27 bp with two dA_2 bulges, to determine the bend induced by the binding of one or two HMG domains, respectively.

MATERIALS AND METHODS

Protein Expression and Purification. HMG-D 112, HMG-D 100, HMG-D 74, HMG-Z 111, and HMG-Z 99 were prepared as previously described (*16*). HMG-D 112 and HMG-Z 111 are the full-length *Drosophila* counterparts of the vertebrate HMG1/2 proteins. HMG-D 100 and HMG-Z 99 are deletion constructs lacking a short stretch of acidic residues at the C-terminus. HMG 74 comprises only the HMG domain of HMG-D 112 and lacks both the acidic region and the basic region (*16*). cHMG1a and cHMG1a 102 were the generous gift of Dr. Wisniewski, Zoologisches Institut, Universität Göttingen, Göttingen, prepared according to Wisniewski and Schulze (*22*). NHP6A was prepared as described by Paull and Johnson (*9*).

Oligonucleotides were synthesized on an Applied Biosystems 394 synthesizer using β-cyanoethyl phosphoramidite chemistry. Fluorescein was conjugated with the oligonucleotides in a final coupling step on the DNA synthesizer using fluorescein phosphoramidite (Applied Biosystems). The succinimidyl ester of 5-carboxytetramethylrhodamine TMRh (Molecular Probes) was attached to the 5′-amino group at the end of a six-carbon linker in dimethyl sulfoxide with 0.2 M carbonate buffer (pH 9.0). Unreacted dyes were removed by chromatography on Sephadex G25 (Pharmacia). The labeled oligonucleotides were purified by reversed-phase HPLC and by 20% polyacrylamide/7 M urea gel electrophoresis.

Complementary fluorescein- and rhodamine-conjugated oligonucleotides were hybridized in 450 mM NaCl, 2 mM $MgCl_2$, and 24 mM sodium citrate, pH 7.0, by slow cooling from 80 to 4 °C. Duplex DNA molecules were purified by 20% native polyacrylamide gel electrophoresis and recovered by electroelution at 4 °C.

The buffer of the DNA samples was changed to the binding buffer [150 mM NaCl, 20 mM HEPES (pH 7.9), and 1 mM EDTA] on a Sephadex G25 column. Absorption measurements were made on this solution from 240 to 650 nm to determine the concentration of the samples and to control the quantity of the labeling; the values indicate 100% labeling.

Spectroscopic Methods. Absorption and fluorescence measurements were taken on a Specord M500 (Zeiss, Germany) and an SLM 8000S or SLM 48000 instrument (SLM Aminco, Urbana, IL). Steady-state fluorescence spectra were corrected for lamp fluctuation and wavelength variations. Polarization artifacts were avoided by using "magic angle" conditions (*23*). The spectra were acquired and processed using the computer program Lab View (National Instruments, Austin, TX). The fluorescence spectra for FRET were collected over a broad range of emission wavelengths (λ_{ex} = 490 nm, λ_{em} = 500−650 nm for the FRET spectrum

and $\lambda_{ex} = 560$ nm, $\lambda_{em} = 570-650$ nm for the acceptor spectrum) and corrected for the buffer signals.

Fluorescence anisotropies were calculated from fluorescence intensity measurements by a vertical excitation polarizer with vertical ($F_{||}$) and horizontal ($F_{\perp}$) emission polarizers according to

$$r = \frac{F_{||} - GF_{\perp}}{F_{||} + 2GF_{\perp}}$$

with the experimental correction factor $G = F_{\perp}/F_{||}$ (*23*).

The fluorescence and absorption spectra of singly labeled molecules with fluorescein (donor) and rhodamine (acceptor) were used to decompose the spectra of doubly labeled samples into donor and acceptor components. All measurements were carried out at 15 °C.

Fluorescence Resonance Energy Transfer (FRET). The rate of resonance energy transfer, k_T, from a fluorescent donor (fluorescein) to an acceptor (TMRh) is related to the donor−acceptor-distance, R, according to

$$k_T = \frac{1}{\tau_D}\left(\frac{R_0}{R}\right)^6$$

with

$$R_0 = 9790(J\kappa^2\phi_D n^{-4})^{1/6}\ \text{Å}$$

τ_D is the fluorescence lifetime of the donor in absence of the acceptor and R_0 the Förster distance depending on the spectral overlap of the dyes, J, the quantum yield of the donor, ϕ_D, the refraction index of the medium, n, and the orientation of the transition dipole moments, κ^2 (*24*−*26*).

For a rapid randomization of the relative donor−acceptor orientation κ^2 is $^2/_3$ (*27*). The low anisotropy of the donor fluorescein implies that this is a good approximation for this study (*28*), so that the efficiency of energy transfer E is sensitive only to the donor−acceptor distance R.

The efficiency of energy transfer E of a donor−acceptor pair at distance R is defined as

$$E = \frac{{R_0}^6}{{R_0}^6 + R^6}$$

Data Analysis. The FRET efficiencies were determined by measuring the intensity of the sensitized emission of the acceptor normalized to the fluorescence of the acceptor alone (*29*). The intensity of the acceptor TMRh emission increases in the presence of FRET, and the intensity of the donor fluorescein emission correspondingly decreases. The spectral dispersions of the fluorescence intensities of the emission spectra $F(\lambda_{em},490)$, where both fluorescein and TMRh absorb, are fitted to the weighted sum of two spectral components: (i) a standard spectrum of donor-only sample and (ii) the fluorescence spectrum of the double-labeled sample excited at 560 nm, where only TMRh absorbs.

$$F(\lambda_{em},490) = aF^D(\lambda_{em},490) + (\text{ratio})_A F(\lambda_{em},560)$$

a and $(\text{ratio})_A$ are the fitted weighting factors of the two spectral components. The fit is made over $\lambda_{em} = 500-540$ nm (where only the donor emits) and $\lambda_{em} = 570-650$ nm (where both the donor and acceptor emit). $(\text{ratio})_A$ is the acceptor fluorescence signal of the FRET measurement normalized by $F(\lambda_{em},560)$ (*28, 29*).

$$(\text{ratio})_A = \frac{F(\lambda_{em},490) - aF^D(\lambda_{em},490)}{F(\lambda_{em},560)} = E\frac{\epsilon^{D,490}}{\epsilon^{A,560}} + \frac{\epsilon^{A,490}}{\epsilon^{A,560}}$$

$(\text{ratio})_A$ is linearly dependent on the efficiency of energy transfer E. It normalizes the measured sensitized FRET signal for the concentration, for the quantum yield of the acceptor, and for any errors in percentage of acceptor labeling. ϵ^D and ϵ^A are the molar absorption coefficients of the donor and acceptor at the given wavelength. $\epsilon^{D,490}/\epsilon^{A,560}$ and $\epsilon^{A,490}/\epsilon^{A,560}$ are determined from the absorbance spectra of doubly labeled DNA molecules and the excitation spectra of singly TMRh-labeled molecules.

Electrophoretic Mobility Shift Assay. The doubly labeled molecules 1xdA_2−DNA and 2xdA_2−DNA (200 nM), with one strand labeled by fluorescein and the other by rhodamine, were incubated during 15 min with different concentrations of HMG-D, HMG-Z, cHMG1a, and their deletion mutants in 20 nM HEPES (pH 7.9), 150 nM NaCl, and 100 μg/mL bovine serum albumin. The samples were loaded on an 8% native polyacrylamide gel and electrophoresed for 40−50 min at 100 V (buffer 0.5× TBE). The gels were directly scanned using a FluorImager 595 (Molecular Dynamics) and analyzed with the ImageQuant software.

DNA Circularization. One dA_2-bulged DNA fragment with complementary ends was used for circularization (GCGCG-TTTTCAATATTTT)·d(GCAAAATATTAAGAAAACGC). First, each oligonucleotide strand was gel purified as described above, and then both were phosphorylated with T4 DNA kinase (one strand of each duplex was 5′-end-labeled with [γ^{32}P]ATP) and finally annealed. The duplexes (∼3 μM) were multimerized with 200 units of T4 DNA ligase (Biolabs) in the commercial T4 DNA ligase buffer (Biolabs) for 1 h at room temperature in the absence or presence of 10 μM HMG-D. The reaction was stopped by the addition of $^1/_{10}$ volume of 10% SDS followed by a phenol/chloroform extraction. The DNA was ethanol precipitated and dissolved in 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA. In some cases the samples were treated by 50 units of exonuclease III, which digests linear but not closed circular DNA molecules, for 15 min at 37 °C and deproteinized by a phenol/chloroform extraction. The samples were loaded onto a 8% native polyacrylamide gel and electrophoresed for 4 h at 15 V/cm in 1× TBE. The gels were dried and autoradiographed.

Molecular Modeling. (A) Modeling of the Dye Position at the DNA Ends. The distribution of dye positions at the DNA helix ends was calculated by taking into account their steric and electronic properties derived from ab initio calculations. A systematic conformational search with force field methods was carried out, resulting in well-defined spatial regions that in principle represent the population of different dye conformers on the DNA helix ends. The majority of the FAM conformers are pointing away from the DNA with an extended linker conformation. This is presumed to be mainly due to the electrostatic repulsion of the 2-fold negatively charged dye and the polyanionic DNA. In contrast, the majority of the zwitterionic TMRh conform-

ers interact with the DNA in the major groove and stack on top of the DNA helix (*30*). The results of these calculations are in agreement with experimental data obtained with anisotropy (*31*) and burst-integrated fluorescence lifetime (BIFL) (*32*). These measurements indicate that TMRh mainly stacks on top of the last base pair (C. Seidel, personal communication), whereas FAM freely rotates in solution. The dye-to-dye distances from a series of seven double-helical DNA molecules determined with FRET (*33*) correlate well with the theoretically calculated distances. These approximate dye positions at the DNA helix ends are used to transform measured FRET efficiencies into DNA structure information.

(B) Modeling of the dA_2−Bulged DNA Molecules. Models of the dA_2−bulged DNA molecules were built by excising two bases from canonical B-type DNA (*34*), reducing the twist of the unopposed bases from 36° to 18°, and reconnecting the backbone of the shorter strand within the program JUMNA (*35*). Thus, in accordance with NMR studies of bulged DNA molecules (*36*−*38*), the unpaired bases were assumed to point inward in a stacked conformation. The angle between the two stems of the single dA_2 bulge was adjusted to match the dye-to-dye distances observed in our FRET experiments. These starting structures were optimized under in vacuo conditions by applying the AMBER 91 force field (*39*) in AMBER 4.1 (*40*). A distance-dependent dielectric constant and a cutoff of 12 Å were used for the electrostatic interactions. The minimization was terminated if the energy gradient of successive steps was below 0.1 kcal/(mol Å).

(C) Modeling of the HMG-D−Bulged DNA Complexes. The first model of the HMG-D−DNA complexes was built on the basis of two NMR structures of HMG proteins but without information on the DNA end-to-end distances from FRET experiments. The 3D structures of HMG-D (ref *6*; PDB entry code 1HMA) and LEF-1 in complex with a 15 bp DNA (ref *13*; PDB entry code 1LEF) were superimposed with their three helices (Ala11-Glu26, Thr33-Arg44, and Ser50-Asp62 in HMG-D fitted to Ala9-Glu24, Ala31-His42, and Ala50-Leu62 in LEF-1). The RMS deviation of the backbone atoms of these residues is 1.0 Å, indicating the high structural similarity of the HMG-D and LEF-1 protein core. Flexible side chains at the surface of HMG-D were adjusted manually to accommodate the bent DNA and to remove unfavorable contacts. HMG-D was docked to the double-helical DNA from the LEF-1 complex. The position of the bulged bases was deduced using the protection pattern obtained from footprinting experiments (D. Payet, unpublished experiments). The dA_2 bulge in the DNA−HMG-D complex was built by excising two nucleotides in the counter strand of the DNA−LEF-1 complex. The bases in the bulge were optimized locally using in vacuo conditions with the AMBER 91 force field (*39*).

Parts of the basic tail of HMG-D (residues 82−87) were modeled using the basic region (residues 77−82) from LEF-1 as template. The loop formed by residues 75−81 was built to fill the gap between the DNA and the N-terminus and the last part of helix 3 of HMG-D (*6*). The loop search algorithm available within SYBYL (*41*) was used to search the binary protein database in this program package for suitable protein fragments. The complex of HMG-D with dA_2−bulged DNA was placed into the center of a box consisting of 28 Na^+ and 10 Cl^- ions and 8929 water molecules using the LEAP program [part of AMBER 4.1 (*40*)]. The AMBER 95 force field (*42*) was applied for a molecular mechanics optimization. A distance-dependent dielectric constant ($\epsilon = 1$) and a cutoff of 12 Å were used for the electrostatic interactions. The minimization was terminated if the energy gradient of successive steps was below 0.1 kcal/(mol Å).

A second model was built when FRET data were available. The region of the DNA duplex that was not in contact with the HMG box was readjusted to match the dye-to-dye distances observed in our FRET experiments. Only small changes had to be introduced to obtain a model with is fully consistent with the FRET data. The kinking angle was altered from 88° to 95° in order to match the experimentally observed dye-to-dye distances. The same protocol as described above was used to energy minimize the second model. The model of the double-bulge DNA in complex with two HMG-D molecules was built by fitting the 5′ end of one single-bulge HMG-D complex onto the 3′ end of a second identical complex. The analysis of the kinking angles was carried out with the program CURVES (*43*). Distances between the averaged dye positions were measured using SYBYL (*41*).

The principal feature of the model is that HMG-D binds on one side of the DNA bulge, with residues Val32 and Thr33 (loop between helix 1 and helix 2) deeply inserted into the "hole" in the minor groove of the bulged DNA, formed by the two missing bases. The overall bend angle of ∼90−100° derives in large part from two large positive roll angles induced by the partial intercalation of Val32 immediately adjacent to the bulge and of Met13 two base steps distant from the Val32 binding site. The conformation of the proximal part of the basic tail in the model is located in the major groove on the opposite face of the DNA to Val32 intercalation and covers some nucleotides at the 5' end of the nonbulged strand region (see Figure 2) that were shown to be protected by footprinting studies (D. Payet, unpublished experiments).

RESULTS

The conformational changes of the DNA induced by binding non-histone chromosomal HMG proteins were measured in solution by fluorescence resonance energy transfer. Figure 3 shows the amino acid sequences of the HMG proteins HMG-D, HMG-Z, and cHMG1a. The first approximately 75 amino acids comprise the DNA-binding HMG domain (box B), followed by a basic region and an acidic tail. The acidic tail contains 12 amino acid residues, of which 10 are either glutamic or aspartic acid. Complexes of these proteins with end-labeled oligonucleotides containing dA_2 bulges were formed. Each single strand of DNA was 5′ labeled with a fluorescent dye, either rhodamine or fluorescein (see Figure 4). In the following measurements we determined the distance in space between the two dyes in the complexes of both the wild type (full-length proteins) and their deletion mutants with the bulged DNA fragments.

HMG1/2 proteins bind without sequence specificity to double-stranded DNA (*44*). To determine the bending angle of the protein−DNA complex, we needed to introduce a specific binding site into the DNA. A suitable such site is provided by a dA_2 bulge (*16*). For our experiments we therefore constructed DNA fragments with one and two defined dA_2 bulge binding sites.

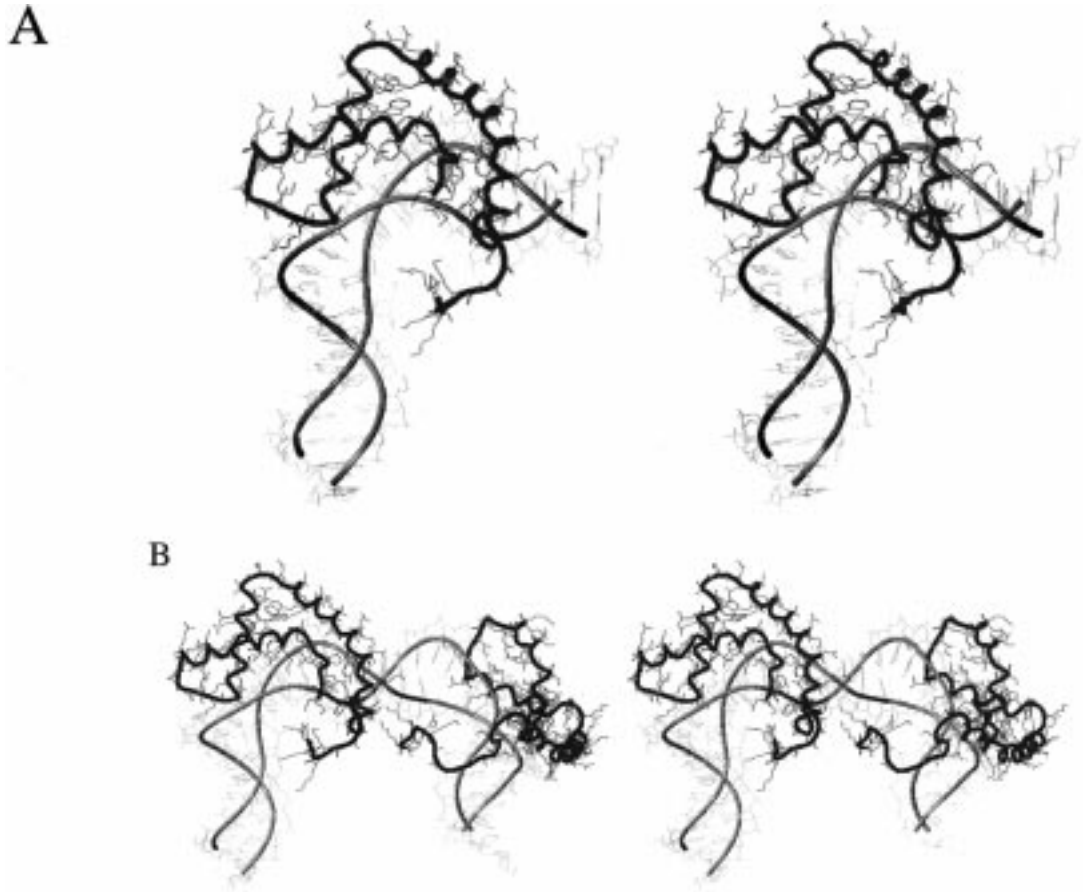

FIGURE 2: Stereoview of the complexes between (dA_2 bulge) DNA and HMG1-like proteins. The HMG domain is from HMG-D. The HMG-box B induced a bend angle in the DNA of 95 ± 5° (A). When two proteins are bound, the overall bend angle is nearly doubled (B). The calculated end-to-end distances in the complex A and B correspond with the measured ones by FRET (Table 2).

```
                1        10        20        30        40        50        60        70        80        90       100       110   116
                |        |         |         |         |         |         |         |         |         |         |         |     |
HMG-D
                  MSDKPKRPLSAYMLWLNSARESIKRENPGIKVTEVAKRGGELWRAM--KDKSEWEAKAAKAKDDYDRAVKEFEANGGSSAANGGGAKKRAKPA--KKVAKKSKKEESDEDDDDESE
HMG-Z
                MSGDRPKRPLSAYMLWLNETREQIKKDNPGSKVTDIAKRGGELWRGL--KDKTEWEQKAIKMKEEYNKAVKEYEANGGTDS---GAPKKRKKAA--AKPAKKAKKKESSEEEEEDESE
cHMG1a
                  MAEKPKRPLSAYMLWLNSARESIKKENPDFKVTEIAKKGGELWRGM--KDKSEWEAKAAKMKEEYEKAMKEFERNGGDKSS-GASTKKRGKAAEKKKPAKKSKKKDSEDDEEEDESD
NHP6A
MVTPREPKKRTTRKKKDPNAPKRALSAYMFFANENRDIVRSENPDITFGQVGKKLGEKWKALTPEEKQPYEAKAQADKKRYESEKELYNATLA

       ++++++++++//////////////////////////////// HMG domain ////////////////////////////////////          ++++++++++++++++++++++----------
```

FIGURE 3: Sequence alignment of HMG proteins HMG-D (*55*, *56*) and HMG-Z (*56*) from *Drosophila*, cHMG1a from *Chironomus* (*22*), and NHP6A (*57*). In the insect proteins the conserved HMG box B is followed by a basic region (+) and an acidic tail (−); in yeast NHP6A the basic region precedes the HMG domain.

dA_2 bulges distort the DNA helix and bend it by about 40−50° (*19*). Our DNA fragments contain a flexible A/T-rich sequence on one side of the bulge and a short dA tract, known to exclude HMG-D binding (*45*) on the other. Footprinting experiments using chemical and enzymatic reagents demonstrated that HMG-D binds preferentially on the side of the bulge containing the flexible DNA sequence (D. Payet, unpublished experiments). Our model of the complex represents this binding property (see Materials and Methods, Figure 2). Correspondingly, this model assumes a head-to-tail binding of two proteins to the DNA with two binding sites. On the basis of the footprinting and the FRET data a head-to-head or a tail-to-tail orientation is very improbable. However, additional experiments are required to determine the protein-binding symmetry to the two binding sites.

The DNA sequences and the position of the fluorescent dyes are shown in Figure 4. Both dyes are covalently bound to the 5′ end of the polynucleotide chain with a six-carbon linker allowing for rotational flexibility. We used an 18 bp DNA fragment with one dA_2 bulge and a 27 bp DNA fragment with two dA_2 bulges to observe the bending induced by one and two HMG proteins, respectively. We verified

single-bulge-DNA ($1xdA_2$)

```
TMRh-5′-CCGAATATTAAGAAAACGGG-3′
      3′-GGCTTATAA--CTTTTGCCC-5′-FAM
```

double-bulge-DNA ($2xdA_2$)

```
TMRh-5′-CCGAATATTAAGAAAGTATTAAGAAAACGGG-3′
      3′-GGCTTATAA--CTTTCATAA--CTTTTGCCC-5′-FAM
```

FIGURE 4: Fluorescently labeled DNA sequences. $1xdA_2$ contains one binding site for HMG proteins (dA_2 bulge) and $2xdA_2$ two binding sites. The fluorescence dyes 6-carboxyfluorescein (FAM) and 5-carboxytetramethylrhodamine (TMRh) were covalently bound to a six-carbon linker at the 5′ end of the oligonucleotides.

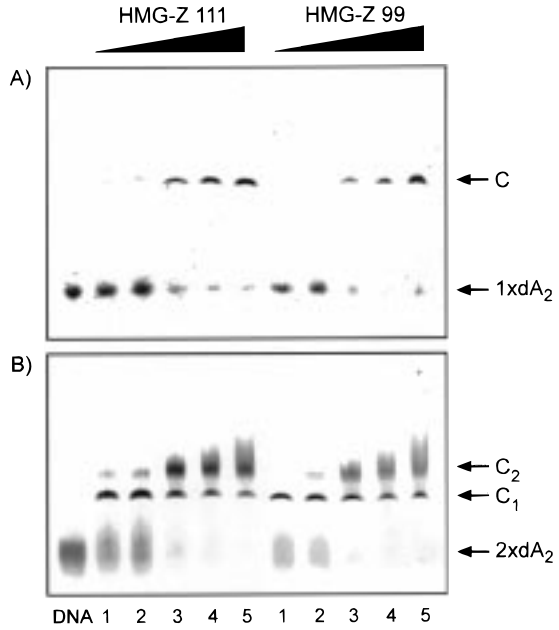

FIGURE 5: Electrophoretic mobility shift assay of HMG-Z 99 and HMG-Z 111 with single-bulge DNA $1xdA_2$ (A) and double-bulge DNA $2xdA_2$ (B). Doubly labeled DNA samples (200 nM) were incubated with an increasing amount of proteins [1 μM (lane 1), 2 μM (lane 2), 5 μM (lane 3), 7.5 μM (lane 4), and 10 μM protein (lane 5)] and separated on 8% native polyacrylamide gels. $1xdA_2$ and $2xdA_2$ indicate the free single- and double-bulge DNA, C is the HMG−$1xdA_2$ complex, and C_1 and C_2 are the 1:1 and 2:1 HMG−$2xdA_2$ complexes.

the binding of one HMG protein to the single-bulge DNA ($1xdA_2$) and two HMG proteins to the double-bulge DNA ($2xdA_2$) by an electrophoretic mobility shift assay in 8% native polyacrylamide gels (see Figure 5).

Spectral Characteristics of the Dye Pair FAM and TMRh. The $\epsilon^{A,490}/\epsilon^{A,560}$ value determined from the excitation spectrum of singly labeled TMRh DNA samples was 0.095 and $\epsilon^{D,490}/\epsilon^{A,560}$ determined from the absorbance spectrum of doubly labeled DNA samples 0.8 (*46*). Both values were used to calculate the efficiency of energy transfer E from the $(\text{ratio})_A$ values. The Förster distance R_0 of the fluorescein−rhodamine pair of the single- and double-bulge molecule was measured to be 50 Å, in agreement with values reported previously (*46*).

FRET Measurements. The decreasing end-to-end distance of DNA by binding HMG proteins can be observed by steady-state fluorescence measurements (Figure 6A). By increasing the amounts of protein, the fluorescence intensity of the acceptor (TMRh) increases while the donor (fluorescein) intensity decreases. Figure 6A shows (i) the normalized fluorescence of donor-only DNA, (ii) free donor−acceptor DNA, and (iii) bound donor−acceptor DNA.

As a control we measured the fluorescence intensity of singly labeled DNA. We observed that the fluorescence intensity of the donor fluorescein is constant in the donor-only sample and independent from the protein concentration (data not shown). Therefore, the donor is not quenched by the proteins. All changes in the donor and acceptor fluorescences are induced by energy transfer due to changes in the DNA conformation and/or a decreasing end-to-end distance. On addition of the protein the anisotropy was measured and found to increase for both dyes as expected. This increase

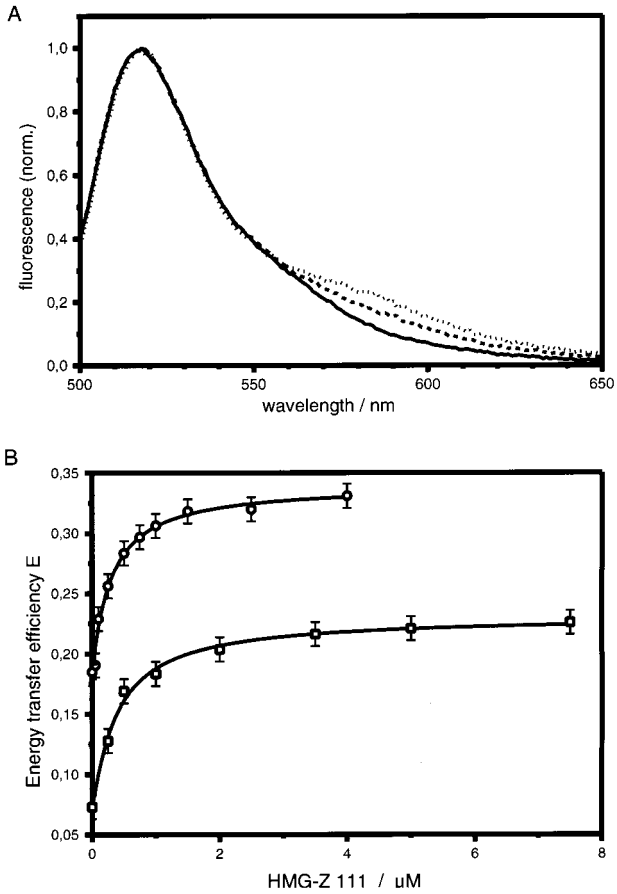

FIGURE 6: (A) Normalized fluorescence spectra of donor-only $1xdA_2$−DNA (—) and donor−acceptor $1xdA_2$−DNA in the absence of protein (- - -) and with 4 μM HMG-Z 111 (···). The DNA concentration was 50 nM. (B) Binding of HMG-Z 111 to $1xdA_2$−DNA (○) and to $2xdA_2$−DNA (□). The data are fitted according to a binding model of one or two ligands binding to DNA with one or two independent binding sites as described by Sevenich et al. (*58*). The binding constants obtained from the fit are 170 nM for a single binding and 290 nM for binding of a second protein at the second binding site. The energy transfer efficiencies for free and bound single-bulge DNA are 0.19 and 0.33; for the double-bulge DNA the energy transfer efficiency increases from 0.07 to 0.23.

is explained by two effects: (i) the increase of mass when the protein binds and (ii) the decrease of the donor lifetime due to the increase of FRET. This also indicates that the binding protein does not interact with the dyes.

Figure 6B shows a typical titration of doubly labeled DNA with HMG proteins. The efficiency of energy transfer E increases from 0.19 to 0.33 by binding one HMG-Z protein to the single-bulge molecule and from 0.07 to 0.23 by binding two proteins to the double-bulge DNA. The end-to-end distances were measured in the complexes with the HMG proteins HMG-D 112, HMG-D 100, HMG-Z 111, HMG-Z 99, cHMG1a, cHMG1a 102, and NHP6A. The affinity of HMG-D 74 for the bulged DNA was too low to permit an accurate determination. The binding constants obtained from the FRET titration plot are in reasonable agreement with those determined by EMSA under slightly different binding conditions.

Table 1 summarizes the measured energy transfer efficiencies of the DNA molecules for all studied HMG proteins. The results show that the full-length proteins as well as the deletion mutants bend the DNA in the same way.

Table 1: Summary of Energy Transfer Distance Measurements in the Absence and Presence of HMG Proteins and Their Deletion Mutants

	single-bulge DNA		double-bulge DNA	
sample	energy transfer	dye distance[a] (Å)	energy transfer	dye distance (Å)
DNA	0.19 ± 0.01	64 ± 0.5	0.07 ± 0.01	76 ± 2
DNA + HMG-D 112	0.32 ± 0.02	56 ± 1	0.24 ± 0.01	60 ± 1
DNA + HMG-D 100	0.33 ± 0.015	56 ± 1	0.26 ± 0.01	59 ± 1
DNA + HMG-Z 111	0.33 ± 0.01	56 ± 0.5	0.23 ± 0.01	61 ± 1
DNA + HMG-Z 99	0.33 ± 0.01	56 ± 0.5	0.25 ± 0.01	60 ± 1
DNA + cHMG1a	0.34 ± 0.01	56 ± 0.5	0.24 ± 0.01	60 ± 1
DNA + cHMG1a 102	0.33 ± 0.01	56 ± 0.5	0.26 ± 0.01	60 ± 1
DNA + NHP6A	0.31 ± 0.015	57 ± 1	nd	nd

[a] The distances were calculated from the energy transfer efficiencies with a Förster distance R_0 of 50 Å (see Materials and Methods, FRET). nd, not determined.

Table 2: Comparison of Measured to Calculated Dye-to-Dye Distances

	dye-to-dye distances (Å)	
	measured with FRET	calculated with molecular modeling
single-bulge DNA (1xdA_2)	64 ± 0.5	64 ± 4
1xdA_2 + HMG	56 ± 0.5	56 ± 4
double-bulge DNA (2xdA_2)	76 ± 2	73 ± 4
2xdA_2 + HMG	60 ± 1	57 ± 4

The FRET efficiency *E* is inversely proportional to the sixth power of the distance between donor and acceptor (see Materials and Methods, FRET) and is thus related to the end-to-end distance of the DNA molecules. The calculated end-to-end distances (see Table 2) take into account the specific location of the dyes with respect to the DNA (*30*). They decrease on binding HMG from 64 to 56 Å in the 18 bp fragment and from 76 to 60 Å in the 27 bp fragment by binding two HMG proteins. In the protein−DNA complexes these distances correspond to overall DNA bend angles of ∼90−100° and ∼180−200°, respectively.

Circularization of DNA Fragments Containing dA$_2$ Bulges. The total calculated bend angle of ∼90−100° for the complexes of the HMG protein with a DNA fragment containing a single dA_2 bulge predicts that the proteins should facilitate the ligation of these fragments to form small microcircles. Using an 18/20 bp dA_2-bulged DNA fragment of essentially the same sequence as that used for the FRET measurements, we observed that under standard circularization conditions HMG-D efficiently promoted the formation of circles containing three repeats of the fragment, i.e., with an average DNA strand length of 57 nucleotides) as well as longer circles (Figure 7). This size is smaller than that of any other circles formed so far by protein-assisted ligation procedures and is fully consistent with a total bend angle of >90° in the complex. Without HMG-D the bulged oligomers also circularized readily, but in this case the preferred circular product of the ligation contained seven bulges. The size of these latter circles is comparable to the optimum size observed by Ulanovsky et al. (*47*) on circularizing 21 bp intrinsically curved DNA fragments. This result thus suggests that the average bend angle for each dA_2 bulge is approximately twice that of each helical repeat of this particular intrinsically curved DNA and is thus consistent with the value of 40−50° interpolated from the data of Gohlke et al. (*19*).

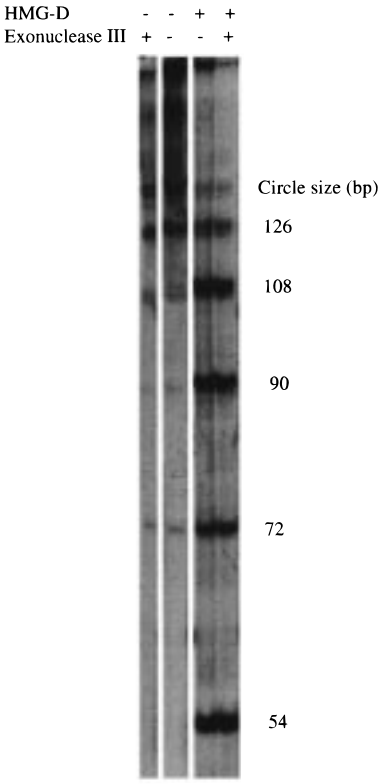

FIGURE 7: Circularization of dA_2−bulged DNA oligomer by HMG-D. The sizes of the circles are indicated as multiples of the short strand (18 bp).

DISCUSSION

The decreasing end-to-end distances in the single-bulge and double-bulge DNA molecules measured by fluorescence resonance energy transfer revealed that the binding of non-histone chromosomal HMG proteins to DNA resulted in a conformational change in the DNA structure consistent with a total bend angle of 95 ± 5°. An independent molecular dynamics simulation (MD) of the HMG-D−DNA complex with the 18 bp DNA fragment with a single dA_2 bulge (1xdA_2) calculated a bending angle very similar (within a few degrees) to the experimentally found 95°. The binding of two proteins to a fragment with two binding sites results in an overall bending angle about twice as large. Our experimentally observed value for the DNA bend angle of 95 ± 5° in the HMG domain−bulged DNA complex agrees to within a few degrees of the value calculated from molecular dynamics simulations of a complex of HMG-D with dA_2-bulged DNA by us but is somewhat greater than the angle of 72 ± 5° similarly calculated by Balaeff et al. (*48*) for a complex of HMG-D with linear DNA in the conformation observed in the TBP−TATA box complex. The 57 nucleotide minimum size of the circles formed with the fragment containing the dA_2 bulge is significantly smaller than that of the 70 bp minimum formed from the corresponding linear fragments (*16*). This is consistent with the bulge making a significant contribution to the overall bending angle in the complex. We envisage that the molecular basis for the recognition of the DNA bulge by protein is high structural complementarity of the loop between helix I and

helix II and the hole in the DNA formed by the two missing bases on one strand. The DNA bend is then induced by the partial intercalation of residues Met13 and Val32 on one side of the bulge coupled with a further flexing of the bulge due to the neutralization of phosphates on the other side of the bulge by the C-terminal basic region of HMG-D binding in the major groove on the inside of the bend.

The other main conclusion of this work is that the bend angle induced by the HMG domain of HMG-D, HMG-Z, and cHMG1 is independent of the C-terminal acidic tail. We also infer that the bend angle is conserved between the homologous HMG domains of HMG-D, HMG-Z, cHMG1a, and NHP6A. The HMG domain of HMG-D and, presumably, by sequence homology, the other domains are structurally homologous to the B-domain of HMG1 (*5*, *6*). The observed bend angle is, however, slightly less than that induced by LEF-1, possibly because HMG-D and its homologues lack the additional minor groove interaction of a tyrosine distal to helix III that is present in LEF-1 (*13*).

A bending angle of about 80−100° for a protein monomer is observed for a large variety of proteins of different sequences and different binding motifs. Also, monomers of the integration host factor IHF (*49*) bend the DNA by about this value (although IHF binds the DNA as a dimer). Thus, local DNA bending by about 80° might be a general value due to structural and/or flexibility limitations of the DNA double helix.

Relevance to Previous Work. In contrast to our measured and calculated bending angle, Heyduk et al. (*1*) reported that a single molecule of cHMG1a induces a bend of about 150°. Using a 30 bp linear DNA sequence, they determined a decrease in the dye-to-dye distance on binding cHMG1a from 100 Å (free, linear DNA) to 50.5 Å (bound DNA) by FRET. This corresponds to an induced bend angle of 150°. We calculate a somewhat larger angle for the binding of two HMG proteins to a 27 bp DNA molecule containing two specific binding sites (dA_2 bulges) but a much smaller angle for the binding of a single HMG protein to a single dA_2 bulge. In the former case, the end-to-end distance of the DNA helix decreases from 76 to 60 Å for all studied HMG proteins. We also observe that the binding of a single cHMG1a molecule to the dA_2 bulge, like the other tested HMG domain proteins, induces an overall bend of ∼95°. We surmise that the bending angle reported by Heyduk et al. is induced by two cHMG1a molecules binding to the DNA fragment rather than the assumed single molecule (the binding of two cHMG1a proteins to the two DNA binding sites was confirmed by electrophoretic mobility shift assays). We note also that a change in direction of the double-helical axis of 150° within the span of the HMG domain binding site would result in steric clashes between the DNA strands on either side of the bend.

Our observation of the high conservation of bend angle induced by the HMG domain in all the HMG protein constructs tested raises the question as to whether the HMG domain by itself bends DNA to the same extent. We know that the HMG domain of HMG-D (HMG-D 74) binds to 75 bp circles with the same affinity as HMG-D 100 containing both the HMG domain and the basic region (*16*). This implies that, when presented with a preformed bend, the DNA contacts made by the HMG domain of HMG-D are independent of the basic region and suggests that this HMG domain has a relatively rigid structure. However, Lnenicek-Allen et al. (*14*) inferred from gel mobility shift assays that the bend angle induced by the HMG domain of human LEF-1 is increased from an estimated ∼57° to ∼77° by a basic C-terminal extension. This extension binds in the major groove on the inside of the DNA bend (*13*) and, like the basic region of HMG-D (*16*), increases the affinity for linear DNA. In our model we observe that basic residues C-terminal to the HMG domain contact the duplex on the opposite side of the bulge to the principal HMG domain contacts and thus could in principle further increase any bend induced by the domain binding on one side of the bulge by flexing the DNA in the region of the bulge itself.

Biological Implications. The average bend angle of ∼95° induced by the HMG domains of the B-class is consistent with the observation that chromosomal HMG proteins containing either a single HMG domain [NHP6A (*9*); HMG-D, S. S. Ner, personal communication] or two HMG domains [HM protein (*50*)] can functionally replace the dimeric HU protein in *Escherichia coli*. Although the precise bend angle induced by HU has not been determined, its close homologue, the IHF heterodimer, bends the DNA duplex by at least 160° in the crystal structure of the protein bound to DNA (*49*); i.e., each monomer bends the DNA by ∼80°. The extent of bending induced by two molecules of HMG-D is thus comparable to that induced by an IHF heterodimer and presumably similar to that induced by HU homo- or heterodimers. This common functionality is, however, achieved by completely different protein−DNA interactions.

HMG-D and its vertebrate counterparts HMG1 and HMG2 have been postulated as structural components of chromatin (*51*−*53*), and HMG1 has been shown to bind to linker DNA on the opposite edge of a positioned nucleosome to histone H1 (*54*), although this latter result is at variance with that reported earlier by Nightingale et al. (*53*). If HMG-D is a bona fide structural component of chromatin, our results imply that it would impose a tight bend on at least part of the linker DNA.

ACKNOWLEDGMENT

We thank Tom Jovin for generous support of our work and Jacek Wisniewski for cHMG1a proteins.

REFERENCES

1. Heyduk, E., Heyduk, T., Claus, P., and Wisniewski, J. R. (1997) *J. Biol. Chem. 272*, 19763−19770.
2. Bustin, M., and Reeves, R. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 54*, 35−100.
3. Bianchi, M. E. (1995) in *DNA-protein: structural interactions* (Lilley, D. M. J., Ed.) pp 177−200, IRL Press, Oxford.
4. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C., and Norman, D. G. (1993) *Nucleic Acids Res. 21*, 3427−3436.
5. Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R. C., Laue, E. D., and Thomas, J. O. (1993) *EMBO J. 12*, 1311−1319.
6. Jones, D. N. M., Searles, M. A., Shaw, G. L., Churchill, M. E. A., Ner, S. S., Keeler, J., Travers, A. A., and Neuhaus, D. (1994) *Structure 2*, 609−627.
7. Pil, P. M., Chow, C. S., and Lippard, S. J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9465−9469.
8. Chow, C. S., Whitehead, J. P., and Lippard, S. J. (1994) *Biochemistry 33*, 15124−15130.
9. Paull, T. T., and Johnson, R. C. (1995) *J. Biol. Chem. 270*, 8744−8754.

10. Giese, K., Cox, J., and Grosschedl, R. (1992) *Cell 69*, 185−195.
11. Giese, K., Pagel, J., and Grosschedl, R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3368−3372.
12. Werner, M. H., Huth, J. R., Gronenborn, A. M., and G. M., C. (1995) *Cell 81*, 705−714.
13. Love, J. J., Xiang, L., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature 376*, 791−795.
14. Lnenicek-Allen, M., Read, C. M., and Crane-Robinson, C. (1996) *Nucleic Acids Res. 24*, 1047−1051.
15. Bianchi, M. E., Beltrame, M., and Paonessa, G. (1989) *Science 243*, 1056−1059.
16. Payet, D., and Travers, A. (1997) *J. Mol. Biol. 266*, 66−75.
17. Pil, P. M., and Lippard, S. J. (1992) *Science 256*, 234−237.
18. Wisniewski, J. R., and Schulze, E. (1994) *J. Biol. Chem. 269*, 10713−10719.
19. Gohlke, C., Murchie, A. I., Lilley, D. M., and Clegg, R. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11660−11664.
20. Parkhurst, K. M., Brenowitz, M., and Parkhurst, L. J. (1996) *Biochemistry 35*, 7459−7465.
21. Ozaki, H., Iwase, N., Sawai, H., Kodama, T., and Kyogoku, Y. (1997) *Biochem. Biophys. Res. Commun. 231*, 553−556.
22. Wisniewski, J. R., and Schulze, E. (1992) *J. Biol. Chem. 267*, 17170−17177.
23. Lakowicz, J. R. (1983) *Principle of fluorescence spectroscopy*, Plenum Press, New York.
24. Clegg, R. M. (1996) Fluorescence Resonance Energy Transfer (FRET), in *Fluorescence imaging spectroscopy and microscopy* (Wang, X. F., and Herman, B., Eds.) pp 179−252, John Wiley & Sons, Inc., New York.
25. Förster, T. (1948) *Ann. Phys. 2*, 55−75.
26. Van der Meer, B. W., Coker, G., III, and Chen, S.-Y. S. (1994) *Resonance energy transfer—Theory and data*, VCH Publischer, Inc., New York.
27. Förster, T. (1949) *Z. Naturforsch. 4a*, 321−327.
28. Clegg, R. M., Murchie, A. I., Zechel, A., Carlberg, C., Diekmann, S., and Lilley, D. M. (1992) *Biochemistry 31*, 4846−4856.
29. Clegg, R. M. (1992) *Methods Enzymol. 211*, 353−388.
30. Hillisch, A. (1998) Ph.D. Thesis, University of Vienna.
31. Vamosi, G., Gohlke, C., and Clegg, R. M. (1996) *Biophys. J. 71*, 972−994.
32. Eggeling, C., Fries, J. R., Brand, L., Gunther, R., and Seidel, C. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 1556−1561.
33. Clegg, R. M., Murchie, A. I., Zechel, A., and Lilley, D. M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 2994−2998.
34. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G., and Ratliff, R. L. (1980) *Nature 283*, 743−745.
35. Lavery, R. (1988) Junctions and Bends in Nucleic Acids: A New Theoretical Modeling Approach, in *DNA Bending & Curvature* (Olson, W. K., Sarma, M. H., Sarma, R. H., and Sundaralingamm, M., Eds.) pp 191−211, Adenine Press, Guilderland, NY.
36. Aboul-ela, F., Murchie, A. I., Homans, S. W., and Lilley, D. M. (1993) *J. Mol. Biol. 229*, 173−188.
37. Rosen, M. A., Live, D., and Patel, D. J. (1992) *Biochemistry 31*, 4004−4014.
38. Rosen, M. A., Shapiro, L., and Patel, D. J. (1992) *Biochemistry 31*, 4015−4026.
39. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) *J. Comput. Chem. 7*, 230−252.
40. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., and Cheatham, T. E., III (1995) *AMBER 4.1*, University of San Francisco.
41. Tripos (1997) *SYBYL 6.4*, Tripos Inc., St. Louis, MO.
42. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
43. Lavery, R., and Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.
44. Bianchi, M. E. (1994) *Mol. Microbiol. 14*, 1−5.
45. 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.
46. Stühmeier, F., Welch, J. B., Murchie, A. I., Lilley, D. M., and Clegg, R. M. (1997) *Biochemistry 36*, 13530−13538.
47. Ulanovsky, L. E., Bodner, M., Trifonov, E. N., and Choder, M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 83*, 862−866.
48. Balaeff, A., Churchill, M. E., and Schulten, K. (1998) *Proteins 30*, 113−135.
49. Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996) *Cell 87*, 1295−1306.
50. Megraw, T. L., and Chae, C. B. (1993) *J. Biol. Chem. 268*, 12758−12763.
51. Ner, S. S., Travers, A. A., and Churchill, M. E. A. (1994) *Trends Biochem. Sci. 19*, 185−187.
52. Travers, A. A., Ner, S. S., and Churchill, M. E. A. (1994) *Cell 77*, 167−169.
53. Nightingale, K., Dimitrov, S., Reeves, R., and Wolffe, A. P. (1996) *EMBO J. 15*, 548−561.
54. An, W., van Holde, K., and Zlatanova, J. (1998) *J. Biol. Chem. 273*, 26289−26291.
55. Wagner, C. R., Hamana, K., and Elgin, S. C. R. (1992) *Mol. Cell. Biol. 12*, 1915−1923.
56. Ner, S. S., Churchill, M. E. A., Searles, M. A., and Travers, A. A. (1993) *Nucleic Acids Res. 21*, 4369−4371.
57. Kolodrubetz, D., and Burgum, A. (1990) *J. Biol. Chem. 265*, 3234−3239.
58. Sevenich, F. W., Langowski, J., Weiss, V., and Rippe, K. (1998) *Nucleic Acids Res. 26*, 1373−1381.

BI990459+