1044–1051 Nucleic Acids Research, 1994, Vol. 22, No. 6

DNA looping by the HMG-box domains of HMG1 and modulation of DNA binding by the acidic C-terminal domain

Michal Štros, Jitka Štokrová[1] and Jean O.Thomas[2,*]
Institute of Biophysics, Czech Academy of Sciences, Královopolská 135, 612 65 Brno, [1]Institute of Molecular Genetics, Czech Academy of Sciences, Flemingovo nám. 2, 166 37 Prague, Czech Republic and [2]Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

Received December 3, 1993; Revised and Accepted February 10, 1994

ABSTRACT

We have compared HMG1 with the product of tryptic removal of its acidic C-terminal domain termed HMG3, which contains two 'HMG-box' DNA-binding domains. (i) HMG3 has a higher affinity for DNA than HMG1. (ii) Both HMG1 and HMG3 supercoil circular DNA in the presence of topoisomerase I. Supercoiling by HMG3 is the same at ~ 50 mM and ~ 150 mM ionic strength, as is its affinity for DNA, whereas supercoiling by HMG1 is less at 150 mM than at 50 mM ionic strength although its affinity for DNA is unchanged, showing that the acidic C-terminal tail represses supercoiling at the higher ionic strength. (iii) Electron microscopy shows that HMG3 at a low protein:DNA input ratio (1:1 w/w; r = 1), and HMG1 at a 6-fold higher ratio, cause looping of relaxed circular DNA at 150 mM ionic strength. Oligomeric protein 'beads' are apparent at the bases of the loops and at cross-overs of DNA duplexes. (iv) HMG3 at high input ratios (r = 6), but not HMG1, causes DNA compaction without distortion of the B-form. The two HMG-box domains of HMG1 are thus capable of manipulating DNA by looping, compaction and changes in topology. The acidic C-tail down-regulates these effects by modulation of the DNA-binding properties.

INTRODUCTION

The high mobility group (HMG) proteins 1 and 2, which are closely related, are amongst the most abundant and ubiquitous non-histone proteins in eukaryotic cell nuclei. Their evolutionary conservation suggests that they are essential for cellular function but their role is not yet clear [1]. They have been implicated in cellular differentiation [2], chromosomal replication [3,4 and references therein], nucleosome assembly [5], and transcription [6–8]. They have also been reported to be associated with transcriptionally inactive (as well as active) sequences [9,10] and to suppress nucleosome assembly *in vitro* [11].

HMG1 and 2 have a tripartite domain structure [12]. The homologous N-terminal (A) and central (B) domains contain the majority of the basic and hydrophobic residues [13] and interact with DNA *in vitro* [14]. Homologous DNA-binding regions (the 'HMG box' [15]) occur in a number of sequence-specific transcription factors and other DNA-binding proteins (reviewed in [16]), some of which have been shown to bend DNA (e.g. the testis determining factor SRY by 85°, and the lymphocyte enhancer-binding factor LEF-1 by 130° [17]). The HMG box is a novel DNA-binding motif, with a distinctive L-shaped structure which has recently been determined by NMR spectroscopy [18,19]. The C-terminal (C) domains of HMG1 and 2 are polyanionic, containing, in the case of bovine HMG1, an unbroken run of 30 glutamic or aspartic acid residues [20, 21]. Binding of calcium to the acidic tail *in vitro* increases the affinity of HMG1 for naked DNA [22], suggesting that interactions involving the tail could, in principle, have profound implications for interactions involving the rest of the molecule in chromatin *in vivo*. In chromatin the acidic C-domain of HMG1 and 2 may interact with the core histones, as demonstrated for the free proteins *in vitro* [14,23]; or with the linker histones, perhaps modulating their interaction with DNA [24, 25]. HMG1 and 2 have been reported to display a preference for single-stranded over double-stranded DNA [26] and to show a preference for synthetic four-way DNA junctions over the corresponding double-stranded DNA [27, 28]. They bind preferentially to supercoiled relative to relaxed DNA, and conserve torsion in negatively supercoiled DNA in the presence of topoisomerase I [29, 30]. They also insert negative supercoils into nicked, or relaxed closed-circular, double-stranded DNA in the presence of DNA ligase, or topoisomerase I, respectively [31, 32].

HMG1 from which the C-tail has been proteolytically removed, leaving the two-domain AB fragment, has been designated HMG3 [33, 25]. We have compared HMG1 and HMG3 in order to determine the effect of the acidic C-terminal domain on DNA-binding and supercoiling by the HMG boxes at low (~ 15 mM)

*To whom correspondence should be addressed

and physiological (~ 150 mM) ionic strength. We have also investigated by electron microscopy the nature of the complexes formed at physiological ionic strength where the differences between HMG1 and HMG3 appear to be greatest. The results show that the acidic C-terminal domain of HMG1 modulates the interaction of the HMG-box A and B domains with DNA and that these domains are able to manipulate DNA structure by looping, compaction and changes in DNA topology. Recent reports, since the completion of this work, show that HMG1 can also bend DNA, as judged by a DNA circularization assay, and this is also a property of the isolated B-domain [34, 35].

MATERIALS AND METHODS

HMG1, HMG3 and plasmid DNA

HMG1 was isolated from calf thymus under non-denaturing conditions and purified as described previously [25], dialysed against 0.14 M NaCl, 10 mM triethanolamine (TEA)/HCl (pH 7.8), 0.5 mM EDTA, 1 mM dithiothreitol (DTT), 0.5 mM phenylmethylsulphonyl fluoride (PMSF) [buffer A; 'physiological' ionic strength (~ 150 mM)] and stored at −70°C. HMG3 (i.e. HMG1 lacking the acidic C-terminal domain) was prepared by limited digestion of HMG1 with trypsin (Serva; TPCK-treated; EC 3.4.21.4) and further purified by ion-exchange chromatography and gel filtration as described [25], and then dialysed and stored as for HMG1. The purity of the proteins was assessed by electrophoresis in SDS/15%-polyacrylamide gels [36] and their concentrations determined as described previously [25]. Amino acid analysis [37] and limited N-terminal sequence analysis [38] of HMG1 and HMG3 were performed as already described.

pBR322 DNA was isolated by alkaline lysis of *E.coli* JRS 856 cells harbouring the plasmid, and the supercoiled form (FI) was purified (twice) in a CsCl gradient, dialysed into buffer A and stored at −70°C.

Nitrocellulose filter-binding assay

pBR322 DNA linearized by *Eco*RI digestion was ^{32}P-labelled by nick-translation and dialysed against several changes of buffer A. After dialysis, DTT and PMSF were added to 1 mM and 0.5 mM, respectively. ^{32}P-Labelled DNA (0.1 μg) was then mixed with various amounts of HMG1 or HMG3 in a final volume of 0.2 ml in either buffer A (final ionic strength ~ 150 mM) or low ionic strength buffer (as buffer A but containing 5 mM NaCl instead of 140 mM NaCl; final ionic strength ~ 15 mM). After 30 min at room temperature, the complexes were diluted two-fold in the same buffers and filtered through prewashed, wet, nitrocellulose filters (0.4 μm); the filters were washed three times with 1 ml buffer, air-dried, and the bound radioactivity counted. The assays were carried out in quadruplicate.

Analysis of supercoiling

Relaxed pBR322 DNA (FI_r) was prepared by treatment of the supercoiled (FI) form with calf thymus topoisomerase I (2 units/μg DNA; prepared in the Institute of Molecular Biology, Bratislava) at 37°C for 90 min either in buffer A or in the same buffer but containing only 40 mM NaCl (ionic strength ~ 50 mM). It was immediately mixed with a second portion of the enzyme and different amounts of HMG1 or HMG3, to give a final DNA concentration of 40−50 μg/ml, and the complexes were incubated for 60 min at 37°C. For analysis of the DNA−topoisomer distribution (but not for electron microscopy; see below), reactions were terminated by addition of 1% SDS and the DNA was isolated by phenol/chloroform extraction. The topoisomer population was analysed by electrophoresis in 1% agarose slab gels in TAE buffer (40 mM Tris base, 20 mM sodium acetate, 1 mM EDTA adjusted with acetic acid to pH 7.2) at 2.8 V/cm for 18 h. The gels were stained with 0.5 μg/ml ethidium bromide for 30 min, destained in water and photographed through a red filter with 254 nm UV-transillumination.

For determination of the change in linking-number of the DNA by HMG3 the topoisomers were resolved by electrophoresis in 1% agarose tube gels containing 50 mM Tris base, 20 mM sodium acetate, 2 mM EDTA and 18 mM NaCl (pH 8.05), with different concentrations (0.02−0.6 μg/ml) of ethidium bromide [39].

Electron microscopy

Complexes containing HMG1 or HMG3 and circular (pBR322) or linear DNA [pBR322 linearized with *Eco*RI or *Pvu*II (both isolated at the Institute of Molecular Biology, Bratislava)] were prepared in buffer A (ionic strength ~ 0.15 M) as described above. Samples were prepared for electron microscopy either by the benzyldimethylalkylammonium chloride (BAC) spreading technique [40] with modifications described previously [41], or by the BAC droplet technique [42] with the following modifications. DNA−protein complexes were fixed in 0.15% (v/v) glutaraldehyde for 24 h at 4°C, or for 15 min at room temperature. The fixed samples (40−50 μg DNA/ml) were diluted 10-fold in buffer A and BAC was added from a stock solution in formamide (2 mg/ml) to a final concentration of 2×10^{-4} % (w/v). After 30 min, droplets of 5 μl were applied to carbon-coated grids and left for 5 min. The grids were then washed with double-distilled water, dehydrated in ethanol, rotary shadowed at 6° with platinum−palladium (80:20, Balzers), viewed with a JEOL JEM 1200EX electron microscope operating at 60 kV, and photographed at a magnification of 10,000× or, in some cases, 30,000×.

Circular dichroism

Circular dichroism measurements were performed using a Jasco J720 spectropolarimeter with a 10 mm pathlength cuvette at 20−24°C. HMG1- and HMG3−DNA complexes with linearized pBR322 DNA were formed as above, and examined at A_{260} = 0.3−0.5 in buffer A containing 1 mM PMSF at 20−24°C. The spectra were analysed using a Jasco J700 program on a Philips PC/AT P3202 computer and expressed as the mean residue ellipticity [θ] (deg. cm^2 dmol^{-1}). Ellipticity values were calculated using an absorption coefficient at 260 nm for the DNA of 6400 l mol^{-1} cm^{-1}.

RESULTS

The aim of this study was to investigate further the effect of the highly acidic C-terminal (C-) domain on the interaction of HMG1 with double-stranded DNA and to examine the nature of the protein−DNA complexes. HMG1 was salt-extracted from calf thymus and the C-domain was removed by limited tryptic digestion under 'structuring conditions'. The remaining two-domain (AB) product, designated HMG3 [33, 25], migrated in an SDS/polyacrylamide gel (Figure 1) as a closely spaced doublet (M_r ~21,400 and ~20,700). It gave a single N-terminal amino acid sequence identical to that of HMG1, indicating that the two

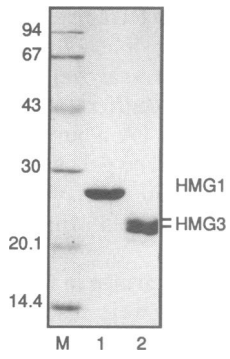

Figure 1. Electrophoretic analysis of HMG1 and HMG3. SDS/15%-polyacrylamide gel electrophoresis of calf thymus HMG1 (lane 1) and HMG3 (lane 2). M: molecular weight markers—phosphorylase b, albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor, and α-lactalbumin (top to bottom). Numbers indicate M_r in kDa.

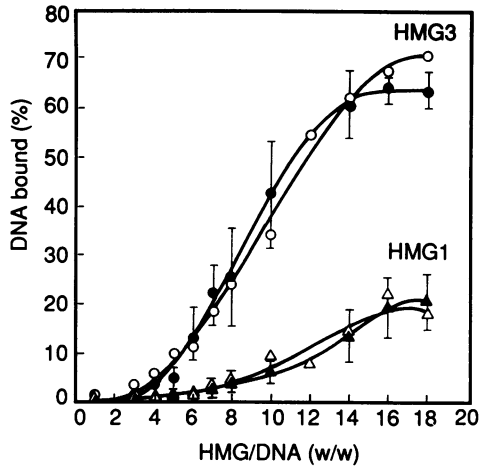

Figure 2. Nitrocellulose filter-binding assay at low and high ionic strengths. ^{32}P-labelled linearized pBR322 DNA was incubated with increasing amounts of HMG1 (▲) or HMG3 (●) in buffer A (i.e. containing 0.14 M NaCl) or in the same buffer but containing only 5 mM NaCl (Δ, ○). The complexes were filtered through nitrocellulose filters and the radioactivity retained was counted. Each curve represents an average of two or three different experiments.

components of HMG3 differed at their C-termini, likely cleavage sites being between residues 179 and 184 which contains five lysine residues (numbering based on the bovine HMG1 sequence [21]). Amino acid analysis (not shown) was consistent with this.

HMG3 has a higher affinity than HMG1 for double-stranded DNA

The effect of removing the acidic C domain from HMG1 on its DNA-binding affinity at low (~ 15 mM) and physiological (~ 150 mM) ionic strength was assessed in a nitrocellulose filter binding assay, using ^{32}P-labelled linearized pBR322 DNA (Figure 2). For HMG1, less than 10% of the DNA was retained up to an input protein:DNA mass ratio (r) of 10:1 (input molar ratio 1130:1), and a maximum of ~ 20% at r = 18. For HMG3, ~ 40% of the DNA was retained on the filters at r = 10 (input molar ratio 1400:1) and ~ 65% at r = 18. Thus there is a significant (~ 4-fold) increase in the affinity of the AB domains of HMG1 for DNA upon removal of the acidic C-domain. For both HMG1 and HMG3, binding was similar at low and high ionic strength.

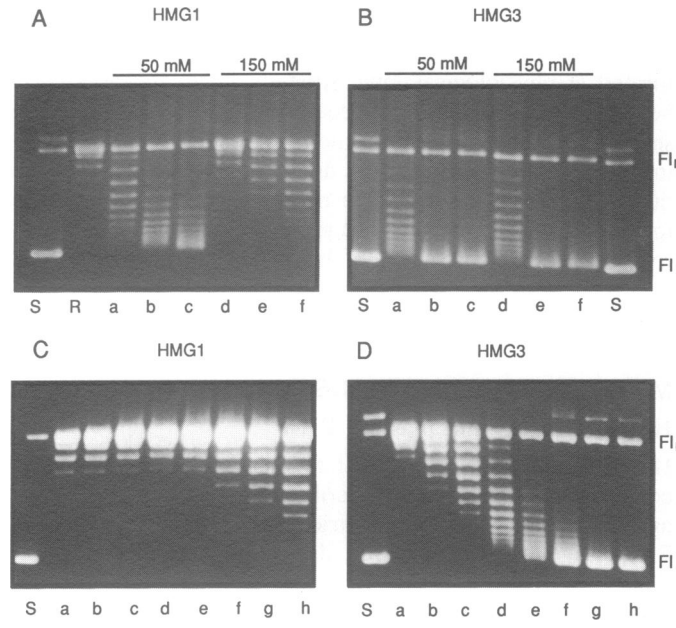

Figure 3. Supercoiling of relaxed circular DNA by HMG1 and HMG3. Circular, relaxed plasmid pBR322 DNA (FI_r) was incubated in the presence of calf thymus topoisomerase I (2 units/μg DNA) with increasing amounts of either HMG1 or HMG3 in either buffer A (0.14 M NaCl, 10 mM TEA/HCl (pH 7.8), 0.5 mM EDTA, 1 mM DTT, 1 mM PMSF; ionic strength ~ 150 mM) or buffer A but containing only 40 mM NaCl (ionic strength ~ 50 mM) for 60 min at 37°C. The deproteinized DNA was analysed in a 1% agarose gel in the absence of ethidium bromide, and the gel stained subsequently. **A** and **B**: supercoiling by HMG1 but not HMG3 is ionic strength-dependent. A, HMG1:DNA (w/w) ratios of 1,3 and 6 at 50 mM (lanes a,b,and c) or 150 mM (lanes d,e and f) ioni strength; B, as A but with HMG3. **C** and **D**: supercoiling at physiological ionic strength (~ 150 mM). C, HMG1:DNA ratios (w/w) of 0, 0.5, 2, 3, 4, 6 and 10 (lanes a–g); D, as C but with HMG3.

DNA supercoiling by HMG1, but not HMG3, is ionic strength dependent

Supercoiling of relaxed, closed circular DNA by HMG1 was much less at physiological than at low ionic strength [11, 43] although, as shown here, there is no effect of ionic strength in this range on the affinity of HMG1 for ds DNA, as judged by nitrocellulose filter binding. To determine whether it is the C-domain in intact HMG1 that represses the ability of the protein to supercoil DNA at physiological ionic strength (~ 150 mM), we compared supercoiling by HMG3 and HMG1 at both low (~ 50 mM) and physiological (~ 150 mM) ionic strengths.

The proteins were incubated with relaxed circular pBR322 (which also contains a small fraction with about four positive and negative supercoils [29 and M.Š., unpublished]) in the presence of calf thymus topoisomerase I at each ionic strength, and the DNA samples were then deproteinized and analysed for supercoiling. Electrophoresis in agarose gels containing different concentrations of ethidium bromide (not shown) showed that both HMG1 and HMG3 at an input HMG:DNA weight ratio of 1 (r = 1) fully suppressed positive supercoils and generated negative supercoils. (Assuming M_r ~25,000 for HMG1, ~ 20,000 for HMG3 and 2.83×10^6 for pBR322 DNA, r = 1 corresponds to 113 molecules HMG1 input per plasmid, or 1 molecule per ~ 39 bp; for HMG3, 140 molecules per plasmid, or 1 molecule

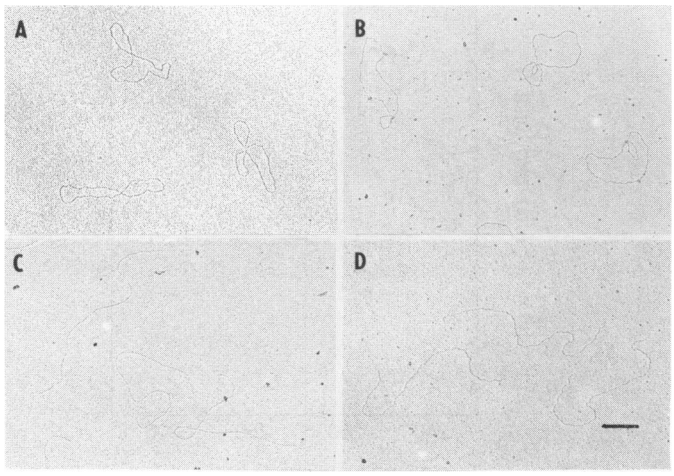

Figure 4. Electron microscopy of HMG1−DNA complexes formed at physiological ionic strength. DNA was either circular, relaxed plasmid pBR322 DNA (FI_r) or the corresponding linearized molecule. **A**, pBR322 DNA relaxed with topoisomerase I (no deproteinization prior to electron microscopy). **B**, Relaxed pBR322 incubated with HMG1 in buffer A (ionic strength ~ 150 mM) at an HMG1:DNA mass ratio (r) of 6 in the presence of topoisomerase I. **C**, pBR322 linearized with *Pvu*II (deproteinized prior to electron microscopy). **D**, Linear pBR322 DNA incubated with HMG1 at r = 6. Bar represents 250 nm.

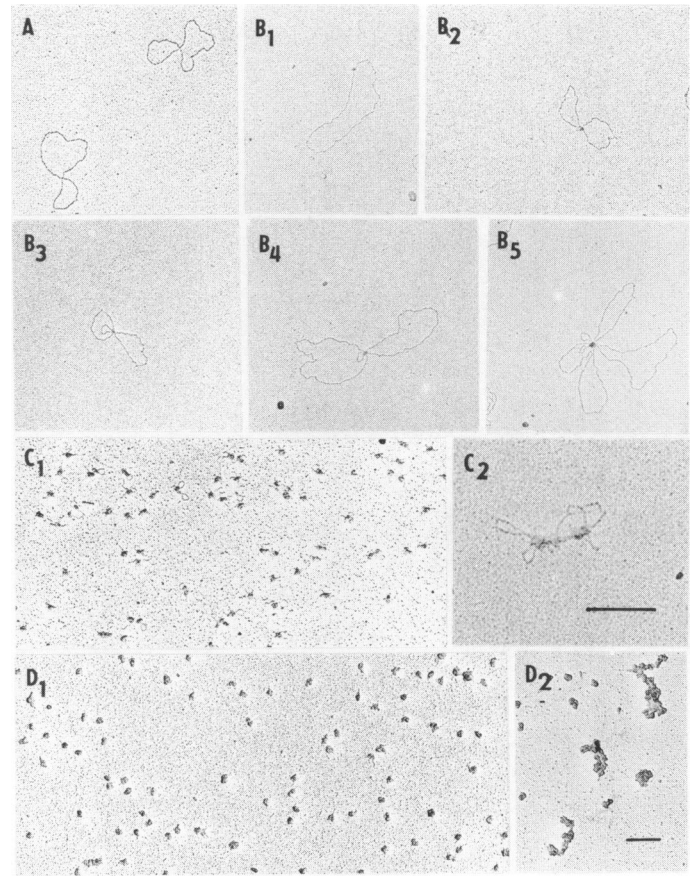

Figure 5. Electron microscopy of HMG3−circular DNA complexes formed at physiological ionic strength. **A**, pBR322 DNA relaxed with topoisomerase I (no deproteinization prior to electron microscopy). $\mathbf{B_1}-\mathbf{B_5}$, Relaxed pBR322 DNA incubated with HMG3 at r = 1 in the presence of topoisomerase I in buffer A. $\mathbf{C_1}-\mathbf{C_2}$, As in B but at r = 3. $\mathbf{D_1}-\mathbf{D_2}$, As in B but at r = 6. Bar represents 250 nm. Magnification in C_2 is twice that in other panels.

per ~ 31 bp.) Electrophoresis in the absence of ethidium bromide (Figure 3A,B) showed that for both HMG1 and HMG3, and at both ionic strengths, the number of supercoils increased with increasing input HMG:DNA ratio. For the same molar input, HMG3 was much more effective at supercoiling than HMG1, consistent with its higher affinity for DNA. However, whereas the effect of HMG3 wasvery similar at both ionic strengths (Figure 3B), at physiological ionic strength supercoiling by HMG1was substantially decreased (Figure 3A) as noted earlier [11, 43], although its DNA-binding affinity was unaffected by ionic strength (Figure 2). Figure 3C and D show supercoiling by HMG1 and HMG3 over a wider protein:DNA input range at physiological ionic strength.

With HMG1, at r = 6, only about 3.4 superhelical turns were generated, corresponding to 1 negative supercoil per ~1280 bp of DNA, with 198 ± 28 HMG1 molecules per molecule of DNA (input, not necessarily bound) responsible for the introduction of one superhelical turn. At an HMG3:DNA mass ratio of 6, the change in linking-number was 17.5 ± 0.8, as determined by the method of Espejo and Lebowitz [39] [no further change at r = 10 (not shown) despite an increase in the amount of HMG3 bound (Figure 2)], corresponding to 1 negative supercoil per ~ 250 bp of DNA. One supercoil thus requires an input of 48 ± 2 HMG3 molecules per DNA molecule, i.e. about 4-fold less than for HMG1 which requires 198 ± 28 molecules, consistent with a ~ 4-fold higher affinity of HMG3 than of HMG1 for DNA at physiological ionic strength, as well as at lower ionic strength (Figure 2).

Although at the lower ionic strength (~ 50 mM) HMG3 still caused supercoiling at a much lower input ratio than HMG1 (c.f. Figure 3A and 3B), the difference between the two was much less than at physiological ionic strength because HMG1 was now relatively much more effective in supercoiling, even though the DNA-binding affinities of the proteins were (probably; Figure 2) unaffected by ionic strength. This suggests that at physiological ionic strength the acidic tail has a role in modulating (down-regulating) the supercoiling ability of the HMG-box domains (A and B), perhaps by engaging in inhibitory interactions with one or other of the domains, or by causing a salt-induced structural change in the protein.

DNA looping, and binding of HMG1 and HMG3 to DNA nodes and crossovers

At an ionic strength of ~150 mM there is a significant difference in the ability of HMG1 and HMG3 to supercoil DNA, which appears to be only partly due to the different affinities of the two proteins for DNA. Complexes formed at this ionic strength were examined by electron microscopy.

Figure 4A shows covalently closed pBR322 DNA after relaxation with topoisomerase I. With HMG1 at r = 6, up to 3−4 beads (10−15 nm diameter) per plasmid were observed (Figure 4B), whereas at r = 1 or 3 there were none (not shown). Parallel analysis of supercoiling (not shown) gave a rough estimate of less than 1 supercoil per bead. In about 75% of the complexes visualized, at least one bead was bound at the site of DNA cross-over or looping (Figure 4B); no beads were

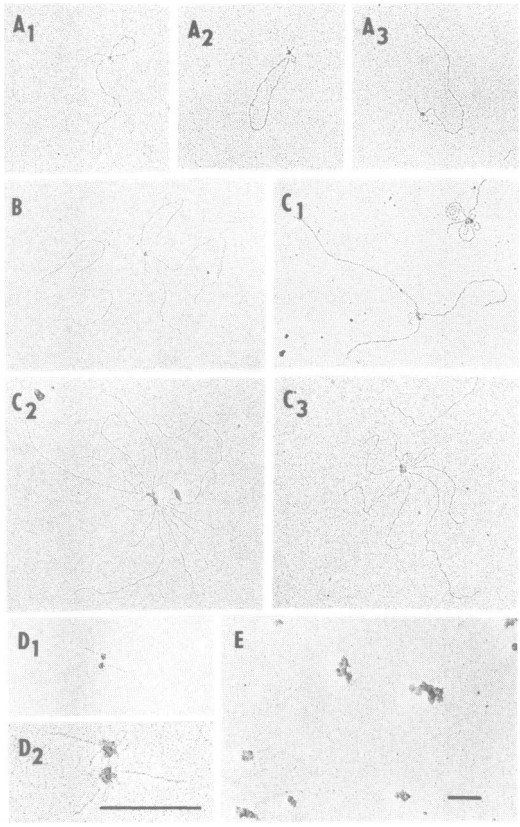

Figure 6. Electron microscopy of HMG3−linear DNA complexes formed at physiological ionic strength. **A**−**C**, Linearized pBR322 DNA (*Pvu*II digest as in Figure 4C) incubated with HMG3 at r = 1 in buffer A. **D**, Linearized pBR322 DNA incubated with HMG3 at r = 3. **E**, as in D but at r = 6. Bar represents 250 nm. Magnification in D2 is three times that in other panels.

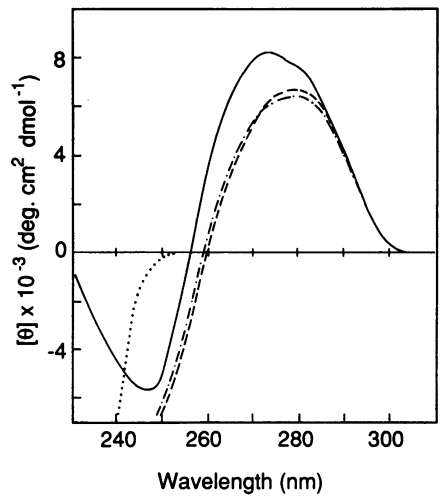

Figure 7. Circular dichroism (CD) spectra of HMG1 and HMG3 complexes with linearized pBR322 DNA at physiological ionic strength (buffer A). Free DNA (—), HMG1−DNA at r = 6 (---), HMG3−DNA at r = 6 (−·−·−·), and HMG3 alone (· · ·).

apparent on linearized pBR322 (Figure 4D). From their size, the beads must contain several HMG1 molecules.

Beads formed at much lower protein:DNA input ratios with HMG3 (Figure 5). Even at r = 1 most (~ 87%) of the complexes visualized with relaxed circular DNA consisted of one HMG3 bead (12−25 nm diameter, and oligomeric) from which emerged several DNA loops of variable size (Figure $5B_2-B_3$). In some cases two or more plasmid molecules were brought together at a bead (Figure $5B_4-B_5$), and in a minor fraction of the complexes there was a (smaller) bead but no DNA-looping (Figure $5B_1$). With linear DNA, looping at the site of HMG3 binding was again observed at r = 1 (Figure $6A_1-A_3$). (In contrast, with HMG1 no looping was observed, even at r = 6.) Similar results were obtained with M13 DNA, suggesting that sequence-specific binding is not involved, although there might be preferential binding to, for example, AT-rich or intrinsically bent regions.

Some complexes contained single DNA molecules with an HMG3 bead bound at the site of looping (58.2% of the 208 complexes scored; Figure $6A_1-A_3$); others contained two or more DNA molecules joined at a bead (41.8% of the complexes scored (Figure 6B and $6C_1-C_3$), respectively. The greater the number of DNA loops the larger the beads, in general. In the absence of HMG3 there were no intermolecular cross-overs, and intramolecular cross-overs occurred only very rarely (Figure 4C). The looped complexes with linear DNA (which were not exposed to topoisomerase 1) argue against the possibility that circular DNA and/or topoisomerase I are a prerequisite for looping; moreover, HMG3 can evidently create DNA nodes or loops in the absence of torsional stress. Topoisomerase I can itself bind to DNA cross-overs and cause looping (e.g. at a 3:1 enzyme:plasmid molar ratio [44]) but under the conditions used here there is no looping of DNA which has been treated with topoisomerase I, fixed, and examined by electron microscopy without deproteinisation (Figure 4A, 5A).

At a higher HMG3:DNA mass ratio (r = 3), the beads were larger and sequestered most of the DNA (Figures $5C_1-C_2$; $6D_1-D_2$). At r = 6 very little, if any, free DNA was apparent (Figure $5D_1$, circular DNA; Figure 6E, linear DNA). The complexes with circular DNA were spherical or somewhat irregular shaped particles (diameter ~ 39−48 nm) (Figure $5D_1$), which sometimes aggregated (Figure $5D_2$). The complexes with linear DNA at the same protein:DNA ratio (r = 6) gave predominantly aggregates (Figure 6E), probably containing several DNA molecules.

Circular dichroism (CD) showed that the DNA in complexes of HMG1 and HMG3 with linearized pBR322 DNA at r = 1 (not shown) or even r = 6 (Figure 7) was typically B-form. The positive band for free DNA at about 274 nm was only slightly suppressed and red-shifted by 4−5 nm. In particular, the absence of a 'Ψ DNA' spectrum [45] indicated that the compacted DNA in the particles generated at r = 6 was not 'condensed' into ordered aggregates, as it is in DNA complexed with H1 and H5 and certain basic peptides [25, 46 and references therein].

DISCUSSION

The effect of the acidic C-domain of HMG1 on DNA binding and supercoiling

The acidic C-terminal domain of HMG1, which contains an unbroken run of 30 glutamic or aspartic acid residues [20, 21], reduces the affinity of the protein for double-stranded DNA ~4-fold at both low (~ 15 mM) and physiological (~ 150 mM) ionic strength. This might result from electrostatic interaction

of the tail with the DNA-binding region(s) in the A and B domains, as first suggested to explain the difference in binding properties between B and BC [44], or alternatively from direct electrostatic repulsion. Both HMG1 and HMG3 cause negative supercoiling of relaxed, closed circular DNA in the presence of topoisomerase I at physiological ionic strength, this evidently being a property of the HMG-box regions. However, because of the differences in affinity for DNA, a higher (~ 4-fold) molar protein:DNA input ratio is required for HMG1, compared with HMG3, to achieve the same supercoiling.

The supercoiling ability of HMG1 at ~ 50 mM ionic strength is greater than at 150 mM, even though the DNA-binding affinity, as judged by nitrocellulose filter-binding, appears unchanged. In contrast, the supercoiling ability of HMG3 appears to be ionic-strength-independent in this range. This suggests some additional negative influence of the C-terminal acidic tail on supercoiling of HMG1 at physiological ionic strength, for example through masking or destabilization of a critical DNA-binding site that is also involved in supercoiling. It may be significant (although the peptides were isolated in strongly denaturing conditions) that the secondary and tertiary structures of the central B domain of HMG1 were lost in the BC peptide consisting of the central and C-terminal domains, as shown by NMR spectroscopy [47, 48], and in contrast to the B-domain, the proteolytically generated BC-didomain gave no supercoiling at physiological ionic strength [47]. Factors that bind to the C-domain, sequestering or shielding the negative charge, might relieve its inhibitory effect on DNA-binding. In the nucleus, where HMG1 is bound to chromatin, such factors might include basic proteins such as histones [14, 23, 25], whose binding to DNA might concomitantly be destabilized, and divalent cations such as Ca^{2+} [22].

A recent study, published after this work had been completed, also showed that removal of the acidic C-domain increased the affinity of the AB didomain for DNA [30]. It was further shown that protection of negative supercoils was greater if the C-tail was present, which was taken to indicate a role for the acidic tail in binding to supercoiled DNA. This would not be incompatible with the indications from our results that the tail is reponsible for the ionic strength dependence of supercoiling by intact HMG1—in other words that it is involved, directly or indirectly, in recognition of supercoiled DNA by the AB domains at physiological ionic strength.

The basis of HMG1-induced supercoiling: DNA looping and binding to cross-overs?

There are, in principle, several possible mechanisms for the change in linking number of closed circular DNA generated by HMG1 and HMG3 in the presence of topoisomerase I. Local strand-separation in the DNA is unlikely to be the cause since HMG1 and HMG2 were found not to lower the melting temperature of natural DNA, poly[d(A-T)] or DNA in H1, H5-stripped nucleosomes [49–51]. Likewise, HMG3 did not lower the melting temperature of either natural DNA or poly[d(A–T)] in either 5 mM NaCl or 140 mM NaCl, in the presence of EDTA (M.Š., unpublished). HMG1 and 2 have been reported to unwind double-stranded DNA in the presence of Mg^{2+} or Ca^{2+} [52], the acidic C-domain of HMG1 being proposed as the 'active site' for the unwinding [53], but as shown here and elsewhere [30], supercoiling does not require the C-domain. Moreover, in Ca^{2+} or Mg^{2+} its charge is presumably neutralized or screened so that its effect is masked [22]. However, as discussed above, the C-domain does apppear to have a role in binding of the AB domains to supercoiled DNA [30] and, as shown here, in the supercoiling of DNA at physiological ionic strength, where it exerts a negative influence.

Wrapping of the DNA around HMG1 or HMG3 might also, in principle, be a cause of the change in linking number of closed circular DNA in the presence of topoisomerase I (c.f. the assembly of nucleosomes on SV40 DNA [54]). However, in an earlier study [32] using acid-extracted HMG1 at an HMG1:SV40 DNA mass ratio of 4:1 (r = 4), in which up to 16 (oligomeric protein) beads of diameter ~ 14.4 nm were observed, less than one (0.5–0.7) negative DNA supercoil per bead was generated [32, 55] compared with one negative supercoil per nucleosome [54]. Moreover, there was no obvious compaction of the DNA [32, 43, 55] and no protection of discrete DNA fragments against micrococcal nuclease digestion [32 and M.Š., unpublished], although such protection would require complexes of some stability. HMG3 (but not HMG1) at r = 6 does cause DNA compaction (Figure 6), but again no discrete DNA fragments were generated upon digestion with micrococcal nuclease (not shown). Circular dichroism showed that in the compact structures the DNA was still in the B-form. A third formal, but perhaps less likely, possibility is that supercoiling arises from a change in the twist of the DNA upon association with HMG1 or HMG3. However this does not appear to be the case for HMG1 [34] nor for HU, an abundant prokaryotic chromosomal protein that binds DNA without sequence specificity and which shows some functional similarities *in vitro* to HMG1 (see below) [56].

Electron microscopy shows that, at r = 1, HMG3 (presumably acting as an oligomer), causes looping of circular and linear DNA at physiological ionic strength and reveals 'beads' at the bases of loops. This is also seen with HMG1 and circular DNA, at a ~ 6-fold higher protein:DNA input ratio, consistent with the lower DNA-binding affinity of HMG1 than of HMG3. The beads are heterogenous in size but on average 10–15 nm in diameter and not dissimilar to the beads observed previously [32, 55], although in that study a 'beads-on-a string' appearance was observed rather than looping. Since the circular DNA used here, produced by treatment of highly negatively supercoiled DNA with calf thymus topoisomerase I, is not completely relaxed (Figure 3) and contains on average four negative and positive supercoils [29; M.Š., unpublished), the cross-overs may serve as preferential binding sites for HMG1 and HMG3. HMG1 and HMG3 generate only negative supercoils so they presumably bind preferentially to cross-overs of the appropriate 'sense'. Binding to, or generation of, cross-overs thus provides a possible mechanism for change in the linking number of DNA by HMG1 and HMG3 in the presence of topoisomerase I.

The preferential localization of HMG3 (and HMG1) beads at the intersections of DNA duplexes, and at the bases of loops, may result from binding to pre-existing cross-overs, or alternatively from interaction with two (or several) DNA sites far apart followed by looping as a result of protein–protein interactions, as occurs with, for example, Sp1 [57, 58]. Looping has also been shown for a number of other proteins including eukaryotic topoisomerases I and II which, like HMG1 and 3, bind preferentially to DNA cross-overs [45]; some type I and II restriction endonucleases [59, and references therein]; and a *Xenopus* protein that was shown subsequently to be HMG1 [60; and references therein].

HMG1: structural rather than sequence preference

Looping through interactions between DNA-bound proteins might suggest some degree of site-specific binding by HMG3, but the variable loop position on linear pBR322 (Figure $6A_1$–A_3)

would seem to argue against this. There is no information on the DNA-binding preference (if any) of HMG3, nor on whether, like HMG1 and 2, it binds preferentially to AT-rich regions [61], of which there are several distinct tracts in pBR322. No strong sequence specificity is evident for HMG1 and 2 , e.g. [62–64].

However, HMG1 and 2 seem able to distinguish between DNA *structures*. They recognise four-way DNA junctions in preference to normal double-stranded DNA [27]. This is not a feature of the acidic C-domain since (recombinant) HMG3 also binds strongly to four-way junctions (S.H.Teo and J.O.T., unpublished results), as do the individual (recombinant) A- and B-domains [28,18]. Indeed the C-domain appears to reduce binding to four-way junctions, when intact HMG1 is compared with truncated recombinant products [28], as it does to double-stranded DNA. HMG1 and 2 bind preferentially to bent regions in linear double-stranded DNA, in particular AT-rich sequences [61], and to kinked (e.g. *cis*-platin modified) DNA [65], and they can also bend DNA [34, 35]. It was shown very recently that HMG1 can substitute for HU, which binds specifically to kinked DNA [66], in facilitating Hin invertasome assembly [34]. This requires the interaction of the Hin and Fis proteins bound at distant sites [56] with looping of the intervening DNA segments, and is facilitated by HMG1-induced DNA bending.

The structures actually recognized in the nucleus by HMG1 might contain bent or kinked DNA (such as occurs, for example, in chromatin-packaged DNA) which has certain features in common with four-way junctions. Alternatively they might be nodes or cross-overs generated by looping, which might resemble four-way junctions [67]. HMG1 would be able to bind to such regions, or perhaps actively create them by looping of the DNA. Electron microscopy (Figures 4–6) shows that the recognition of cross-overs, like the ability to bind to four-way junctions (S.H.Teo and J.O.T., unpublished), is a property of the HMG-box domains of HMG1, and does not require the acidic tail. The recognition of synthetic four-way junctions by HMG1 may therefore be a reflection of a more general function of HMG1 within the cell nucleus, namely manipulation of DNA structure by, for example, looping or, at higher concentrations, compaction/packaging, a function which is manifest primarily through the two HMG-box domains.

ACKNOWLEDGEMENTS

We thank Juliette Jacoby for amino acid analysis, Dr Len Packman for protein sequence determination; Dana Formánková, Dagmar Zachová and Christine Rees for skilful assistance in parts of this work; and Colin Hardman and Teo Soo Hwang for helpful comments on the manuscript. We thank the European Molecular Biology Organization for the award to M.Š.of a Short-Term EMBO Fellowship to work in Cambridge, and the SERC and The Wellcome Trust for support (J.O.T.). M.Š. was also supported by a Czechoslovak Academy of Sciences Operating Grant (50410).

REFERENCES

1. Bustin, M., Lehn, D.A. and Landsman, D. (1990) *Biochim. Biophys. Acta.* **1049**, 231–243.
2. Seyedin, S.M., Pehrson, J.R. and Cole, R.D. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5988–5992.
3. Bonne, C., Sautière, P., Duguet, M and DeRecondo, A.M. (1982) *J. Biol. Chem.* **257**, 2722–2725.
4. Alexandrova, E.A. and Beltchev, B.G. (1988) *Biochem. Biophys. Res. Commun.* **154**, 918–927.
5. Bonne-Andrea, C., Harper, F., Sobczak, J. and DeRecondo, A.-M. (1984) *EMBO J.* **3**, 1193–1199.
6. Stoute, J.A. and Marzluff, W.F. (1982) *Biochem. Biophys. Res. Commun.* **107**, 1279–1284.
7. Tremethick, D.J. and Molloy, P. (1988) *Nucleic Acids Res.* **16**, 11107–11123.
8. Singh, J. and Dixon, G.H. (1990) *Biochemistry* **29**, 6295–6302.
9. Blanco, J., States, J.C. and Dixon, G.H. (1985) *Biochemistry* **24**, 8021–8028.
10. Postnikov, Yu.V., Shick, V.V., Belyavsky, A.V., Khrapko, K.R., Brodolin, K.L., Nicolskaya, T.A. and Mirzabekov, A.D. (1991) *Nucleic Acids Res.* **19**, 717–725.
11. Waga, S., Mizuno, S., Yoshida, M. (1989) *Biochim. Biophys. Acta* **1007**, 209–214.
12. Reeck, G.R., Isackson, P.J. and Teller, D.C. (1982) *Nature* **300**, 76–78.
13. Cary, P.D., Turner, C.H., Mayes, E. and Crane-Robinson, C. (1983) *Eur. J. Biochem.* **131**, 367–374.
14. Carballo, M., Puigdomènech, P. and Palau, J. (1983) *EMBO J.* **2**, 1757–1764.
15. Jantzen, H.-M., Admon, A., Bell, S.P. and Tjian, R. (1990) *Nature* **344**, 830–836.
16. Ner, S. (1992) *Current Biology* **2**, 208–210.
17. Giese, K., Cox, J. and Grosschedl, R. (1992) *Cell* **69**, 185–195.
18. Weir, H.M., Kraulis, P.J., Hill, C.S., Raine, A.C.R., Laue, E.D. and Thomas, J.O. (1993) *EMBO J.* **12**, 1311–1319.
19. Read, C.M., Cary, P.D., Crane-Robinson, C., Driscoll, P.C. and Norman, D.G. (1993) *Nucleic Acids Res.* **21**, 3427–3436
20. Kaplan, D.J. and Duncan, C.H. (1988) *Nucleic Acids Res.* **16**, 10375.
21. Walker, J.M., Hastings, J.R.B. and Johns, E.W. (1978) *Nature* **271**, 281–282.
22. Štros, M., Bernués, J. and Querol, E. (1990) *Biochem. Int.* **21**, 891–899.
23. Bernués, J., Espel, E. and Querol, E. (1986) *Biochim. Biophys. Acta* **866**, 242–251.
24. Kohlstaedt, L.A., Sung, E.,C., Fujishige, A. and Cole, R.D. (1987) *J. Biol. Chem.* **262**, 524- 526.
25. Štros, M. and Vorlíčková, M. (1990) *Int. J. Biol. Macromol.* **12**, 282–288.
26. Isackson, P.J., Fishback, J.L., Bidney, D.L. and Reeck, G.R. (1979) *J. Biol. Chem.* **254**, 5569–5572.
27. Bianchi, M.E., Beltrame, M. and Paonessa, G. (1989) *Science* **243**, 1056–1059.
28. Bianchi, M.E., Falciola, L., Ferrari, S and Lilley, D.M.J. (1992) *EMBO J.* **11**, 1055–1063.
29. Sheflin, L.G. and Spaulding, S.W. (1989) *Biochemistry* **28**, 5658–5664.
30. Sheflin, L.G., Fucile, N.W. and Spaulding, S.W. (1993) *Biochemistry* **32**, 3238–3248
31. Javaherian, K., Liu, L.F. and Wang, J.C. (1978) *Science (Wash.)* **199**, 1345–1346.
32. Mathis, D.J., Kindelis, A. and Spadafora, C. (1980) *Nucleic Acids Res.* **8**, 2577–2590.
33. Goodwin, G.H., Walker, J.M. and Johns, E.W. (1978) *Biochim. Biophys. Acta* **519**, 233–242.
34. Paull, T.T., Haykinson, M.J. and Johnson, R.C. (1993) *Genes Dev.* **7**, 1521–1534.
35. Pil, P.M., Chow, C.S. and Lippard, S.J. (1993). *Proc. Natl Acad. Sci. USA* **90**, 9465–9469.
36. Laemmli, U.K. (1970) *Nature* **227**, 680–685.
37. Clark, D.J. and Thomas, J.O. (1986) *J. Mol. Biol.* **187**, 569–580.
38. Hill, C.S. and Thomas, J.O. (1990) *Eur. J. Biochem.* **187** 145–153.
39. Espejo, R. and Lebowitz, J. (1976) *Anal. Biochem.* **72**, 95–103.
40. Portmann, R., Schaffner, W. and Birnstiel, M. (1976) *Nature* **264**, 31–34.
41. Štokrová, J., Vojtíšková, M. and Paleček, E. (1989) *J. Biomol. Struct. Dynam.* **6**, 891–898.
42. Thoma, F., Koller, T. and Klug, A. (1979) *J. Cell. Biol.* **83**, 403–427.
43. Duguet, M., Bonne, C. and de Recondo, A.-M. (1981) *Biochemistry* **20**, 3598–3603.
44. Zechiedrich, E.L. and Osheroff, N. (1990) *EMBO J.* **9**, 4555–4562.
45. Fasman, G.D., Schaffhausen, B., Goldsmith, L. and Adler, A. (1970) *Biochemistry* **9** 2814–2822.
46. Hill, C.S., Martin, S.R. and Thomas, J.O. (1989) *EMBO J.* **8**, 2591–2599.
47. Carballo, M., Puigdomènech, P., Tancredi, T. and Palau, J. (1984) *EMBO J.* **3**, 1255–1261.
48. Cary, P.D., Turner, C.H., Leung, I., Mayes, E. and Crane-Robinson, C. (1984) *Eur. J. Biochem.* **143**, 323–330.
49. Marekov, L.N., Beltchev, B.G. and Pivec, L. (1984a) *Biochem. Biophys. Res. Commun.* **120**, 782–788.
50. Butler, A.P., Mardian, J.K.W. and Olins, D.E. (1985) *J. Biol. Chem.* **260**, 10613–10620.

51. Štros, M. and Kleinwächter, V. (1987) *Biochem. Biophys. Acta* **910**, 163–170.
52. Makiguchi, K., Chida, Y., Yoshida, M. and Shimura, K. (1984) *J. Biochem.* **95**, 423–429.
53. Yoshida, M. (1987) *J. Biochem.* **101**, 175–180.
54. Germond, J.E., Hirt, B., Oudet, P., Gross-Bellard, M. and Chambon, P. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1843–1847.
55. Bonne, C., Duguet, M. and DeRecondo, A.-M. (1980) *Nucleic Acids Res.* **8**, 4955–4968.
56. Haykinson, M.J. and Johnson, R.C. (1993) *EMBO J.* **12**, 2503–2512.
57. Su, W., Jackson, S., Tjian, R and Echols, H. (1991) *Genes Dev.* **5**, 820–826.
58. Mastrangelo, I.A., Courey, A.J., Wall, J.S., Jackson, S.P. and Hough, P.V.C. (1991) *Proc. Natl. Acad. Sci. USA* 88, 5670–5674.
59. Topal, M.D., Thresher, R.J., Conrad, M. and Griffith, J. (1991) *Biochemistry* **30**, 2006–2010
60. Ghrir, R., Mignotte, B. and Guéride, M.(1991) *Biochimie* **73**, 615–617.
61. Brown, J.W. and Anderson, J.A. (1986) *J. Biol. Chem.* **261**, 1349–1354.
62. Shooter, K.V., Goodwin, G.H. and Johns, E.W. (1974) *Eur. J. Biochem.* **47**, 263–270.
63. Goodwin, G.H., Shooter, K.V. and Johns, E.W. (1975) *Eur. J. Biochem.* **54**, 427–433.
64. Watt, F and Molloy, P.L. (1988) *Nucleic Acids Res.* **17**, 1197–1214
65. Pil, P.M. and Lippard, S.J. (1992) *Science* **256**, 234–236
66. Pontiggia, A., Negri, A, Beltrame, M. and Bianchi, M.E. (1993) *Mol. Microbiol.* **7**, 343–350.
67. Lilley, D.M.J. (1992) *Nature* **357**, 282–283.