



**HAL**  
open science

## Determinants of Specific Binding of HMGB1 Protein to Hemicatenated DNA Loops

Sandrine Jaouen, Leanne de Koning, Claire Gaillard, Eva Muselíková-Polanská, Michal Štros, François Strauss

► **To cite this version:**

Sandrine Jaouen, Leanne de Koning, Claire Gaillard, Eva Muselíková-Polanská, Michal Štros, et al.. Determinants of Specific Binding of HMGB1 Protein to Hemicatenated DNA Loops. *Journal of Molecular Biology*, 2005, 353, pp.822-837. 10.1016/j.jmb.2005.08.073 . hal-00009564

**HAL Id: hal-00009564**

**<https://hal.science/hal-00009564>**

Submitted on 13 Oct 2005

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Determinants of specific binding of HMGB1 protein to hemicatenated DNA loops**

**Sandrine Jaouen<sup>1</sup>, Leanne de Koning<sup>1</sup>, Claire Gaillard<sup>1</sup>, Eva Muselíková-Polanská<sup>2</sup>, Michal Štros<sup>2,\*</sup>, and François Strauss<sup>1,\*</sup>**

<sup>1</sup>Institut Jacques Monod, 2 place Jussieu, F-75251 Paris 05, France

<sup>2</sup>Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, CZ-612 65 Brno, Czech Republic

\* *Corresponding authors:* e-mail addresses: strauss@ijm.jussieu.fr, stros@ibp.cz

*Abbreviations used:* HMG, high mobility group; hcDNA, hemicatenated DNA loops; 4WJ, four-way DNA junction; IPTG, isopropyl- $\beta$ ,D-thiogalactopyranoside

## Summary

Protein HMGB1 has long been known as one of the most abundant non-histone proteins in the nucleus of mammalian cells, and has regained interest recently for its function as an extracellular cytokine. As a DNA-binding protein, HMGB1 facilitates DNA-protein interactions by increasing the flexibility of the double helix, and binds specifically to distorted DNA structures. We have previously observed that HMGB1 binds with extremely high affinity to a novel DNA structure, hemicatenated DNA loops (hcDNA), in which double-stranded DNA fragments containing a tract of poly(CA)·poly(TG) form a loop maintained at its base by a hemicatenane. In this paper we show that the single HMGB1 domains A and B, the HMG-box domain of sex determination factor SRY, as well as the prokaryotic HMGB1-like protein HU, specifically interact with hcDNA ( $K_d \sim 0.5$  nM). However, the affinity of full-length HMGB1 for hcDNA is three orders of magnitude higher ( $K_d < 0.5$  pM) and requires the simultaneous presence of both HMG-box domains A and B plus the acidic C-terminal tail on the molecule. Interestingly, the high affinity of the full-length protein for hcDNA does not decrease in the presence of magnesium. Experiments including a comparison of HMGB1 binding to hcDNA and to minicircles containing the CA/TG sequence, binding studies with HMGB1 mutated at intercalating amino-acid residues (involved in recognition of distorted DNA structures), and exonuclease III footprinting, strongly suggest that the hemicatenane, not the DNA loop, is the main determinant of the affinity of HMGB1 for hcDNA. Experiments with supercoiled CA/TG-minicircles did not reveal any involvement of left-handed Z-DNA in HMGB1 binding. Our results point to a tight structural fit between HMGB1 and DNA hemicatenanes under physiological conditions, and suggest that one of the nuclear functions of HMGB1 could be linked to the possible presence of hemicatenanes in the cell.

*Keywords:* alternative DNA conformations; DNA knot and hemicatenane; DNA topoisomers fractionation in chloroquine gels; high-mobility group proteins; HMG-box domain

## Introduction

The nuclear high-mobility group protein HMGB1 (formerly known as HMG-1<sup>1</sup>) is a highly abundant protein in the nuclei of mammalian cells, present in about one molecule per 10-20 nucleosomes. HMGB1 contains two copies of a polypeptide domain of about 80 amino-acids, the HMG-box motif, which is found in a large number of proteins collectively referred to as the HMG-box family of proteins<sup>1-3</sup>. Despite the fact that sequence conservation of the HMG-box domain between members of the family is not strict, the HMG-box domains are generally considered as sharing a common function linked to the characteristics of their interactions with DNA<sup>4,5</sup>. This is supported by several lines of evidence: first, the HMG-box family includes many transcription factors and factors involved in the control of development or differentiation<sup>3,4</sup>; second, all HMG-box domains studied so far have shown an affinity for DNA<sup>6-8</sup>; and third, numerous studies have shown a modulation of DNA-protein interactions by HMGB1, as this protein can facilitate the assembly of complexes involved in recombination and transcription (review by Bustin<sup>9</sup>, see also Bonaldi *et al.*<sup>10</sup>) as well as nucleosome remodelling<sup>11</sup>. The intracellular role of HMGB1 in association with chromatin does not preclude other functions, however, such as an extracellular role in mediation of inflammation mechanisms, tumor growth and metastasis<sup>12,13</sup>; recent review<sup>14</sup>).

HMGB1, as well as many other proteins containing HMG-box domains, can not only exhibit high affinity to distorted DNA conformations such as supercoiled DNA, four-way DNA junction, DNA minicircles, cisplatin-modified DNA and DNA bulges (for reviews see<sup>6,7,8,15</sup>), they can also actively distort DNA by bending or looping<sup>16-20</sup> or changing DNA-topology<sup>18,21,22</sup>. We have recently reported that protein HMGB1 also binds with unusually high affinity to a novel DNA structure, hemicatenated DNA loops (hcDNA)<sup>23-26</sup>, obtained by reassociation of the strands of DNA fragments containing a tract of the CA-microsatellite poly(CA)·poly(TG). In this structure, double-stranded DNA forms a loop maintained at its base by a hemicatenane, i.e. a junction between two DNA duplexes in which one of the strands of one duplex passes between the two strands of the other duplex, and reciprocally (Figure 1). The novelty of this structure, the strength of its interactions with HMGB1, as well as recent results showing that tumor suppressor p53 also binds to hcDNA with high affinity<sup>27</sup> have encouraged us to further study the interactions of hcDNA with HMGB1 and with HMG-box protein domains.

The present work was aimed at understanding the high-affinity binding of HMGB1 to hcDNA by addressing three questions. First, is it the HMG-box domain itself that is responsible for the high affinity of interaction with hcDNA as is known to be the case with other DNA structures such as four-way junctions or small DNA circles? Second, is the high affinity of HMGB1 for hemicatenated DNA loops due to the presence of a hemicatenane or to distortions of DNA resulting from the

presence of a small loop in the structure? And third, what are the characteristics of the interactions of HMG-box domains with hcDNA as compared to their interactions with other distorted DNA substrates?

## Results

### Binding specificity of HMGB1 and its domains for hcDNA

The fundamental role of HMG-box domains in the specific recognition of distorted DNA structures by proteins of the HMG-box family being well established, we first set out to verify whether this was also true with hcDNA and if the individual domains showed the same binding characteristics as the full-length proteins. Binding experiments were thus carried out with the domains represented schematically in Figure 2(a), which were expressed in *Escherichia coli* as (His)<sub>6</sub>-fusion proteins and purified by chromatography (alternatively the AB-didomain was also prepared by tryptic cleavage of the mammalian protein; see Materials and Methods for detailed information on domain expression and purification). Figure 2(b) presents a polyacrylamide-SDS gel of the purified domains, showing the high level of purity that was obtained after FPLC purification. As could be anticipated, the AB didomain and both individual A and B domains showed a clear specificity toward hcDNA. On the polyacrylamide gels shown in Figure 3 they bound all of hcDNA while showing no complex formation with the linear DNA present in the samples. This was observed in interactions of hcDNA with limiting amounts of the proteins (Figure 3, left panels), as well as in competition experiments with unlabelled bacteriophage lambda DNA (Figure 3, right panels). The domain B having only the C-terminal extension, residues 92-180, gave the same result as the domain B extending on both sides; in contrast, neither the “short” domain B (lacking both N- and C-terminal extensions) nor the acidic C-terminal domain gave rise to shifted bands (data not shown). Therefore, as is the case with the other distorted DNA structures, specific binding to hcDNA is again a property of the HMG-box domain.

### Comparison of binding affinities of HMGB1, domains, and mutants: highest affinity of full-length HMGB1

While sharing hcDNA binding specificity with its HMG-box domains, the full-length HMGB1 protein displayed several unique binding characteristics. First, binding of the full-length protein resulted in only two retarded bands having well-defined mobilities (Figure 3(a), bands labeled C1 and C2), intermediate retarded bands being never observed upon decreasing the protein concentration or increasing the competitor DNA concentration. In contrast, the A and B single domains formed an almost continuous series of many retarded bands (Figure 3(c) and 3(d), left panels), the decrease in mobility of hcDNA being very small at low protein concentration while becoming equivalent to the shift of full-length HMGB1 when the concentration of the single domains was high, reflecting the gradual fixation of several molecules of the domains. Second, the difference between the full-length protein and its domains was even more striking when hcDNA

binding was performed in the presence of competitor DNA (Figure 3, right panels). Complexes C1 and C2 of HMGB1 with hcDNA were almost completely resistant to the presence of competitor (Figure 3(a), right panel), only a slight decrease in band C2 being seen at the extremely high competitor DNA concentration of 400  $\mu\text{g/ml}$  corresponding to a non-specific to specific DNA weight ratio of  $10^5$ . In contrast, complexes formed with the B domain were strongly decreased by a competitor DNA concentration of 40  $\mu\text{g/ml}$  (non-specific to specific DNA ratio =  $10^4$ ) (Figure 3(d) right) while complexes with the A domain had almost disappeared at the same concentration (Figure 3(c) right).

These observations reflect a strong difference in affinity for hcDNA between HMGB1 and its domains. Assuming a simple binding model, gel retardation experiments allow for a quick estimate of binding constants using the formula  $K_d = [P] - [D]/2$ , which relates the dissociation constant  $K_d$  to the total protein and DNA concentrations at 50% binding, respectively [P] and [D], this formula becoming  $K_d \approx [P]$  when  $[D] \ll [P]$ , i.e. when the DNA fragment concentration is significantly lower than  $K_d$ . The dissociation constant of full-length HMGB1 could not be directly measured here since the protein concentration at 50% binding in Figure 3(a) was similar to the DNA fragment concentration (2 pM), indicating a  $K_d \ll 2$  pM. (Indeed, previous gel retardation experiments using DNA concentrations as low as was possible with the detection techniques used<sup>24</sup> have shown a  $K_d$  lower than 0.16 pM for HMGB1 binding to hcDNA, i.e. an affinity constant at least equal to the unusually high value of  $5 \cdot 10^{12}$ ; in accordance the hcDNA-HMGB1 complex was found to be extremely stable, with a half-life of 70 min). In contrast, with A and B domains the protein concentration at 50% binding was more than 10 fold higher than the DNA concentration (Figs. 3C, 3D), yielding  $K_d$  estimates of  $\sim 0.5$  nM for both single domains. This similarity in binding affinities of A and B domains was no longer true with the domain mutants obtained by alanine mutagenesis of the DNA-intercalating residues, i.e. the residues that intercalate between base pairs in the complex and have been shown to be important for interactions of HMG-box domains with curved or distorted DNA<sup>28-31</sup>. As shown in Figure 3(c), replacement of phenylalanine at position 38 within domain A resulted in a tenfold lower affinity for hcDNA, with a  $K_d$  of  $\sim 5$  nM, and its binding to hcDNA was completely abolished by a  $10^3$ -fold excess of  $\lambda$  DNA (note that replacement of Phe38 for Ala had no significant effect on binding of the HMGB1 domain A to linear DNA<sup>29</sup>, M.Š., paper in preparation). In contrast, mutations of intercalating residues Phe103 or Ile122 or even the double mutation Phe103+Ile122 of the domain B did not show a major decrease in hcDNA-binding affinities as compared to the wild-type domain (Figure 3(d)).

The results with the AB didomain were somewhat unexpected, as three bands could be seen on the gels (Figure 3(b)). The two upper bands looked very similar to bands C1 and C2 of HMGB1, their

slightly lower mobility being probably due to the charge difference resulting from the removal of the acidic C-terminal domain. The lower band 3, which has never been observed with full-length HMGB1, appeared only at low protein concentration or high competitor amounts and was consistently observed with the recombinant AB didomain as well as with the didomain obtained by tryptic cleavage of mammalian HMGB1, as on the gel shown here. Whereas the didomain usually shows a higher affinity than the complete protein for most distorted DNA structures, this was clearly not the case here. Binding experiments performed at hcDNA concentrations varying between 1 nM and 1 pM showed an affinity of the AB didomain that was intermediate between those of the single domains and of the full-length protein, with a  $K_d$  of ~50 pM (data not shown). In addition, linear DNA was a clear competitor of AB didomain binding to hcDNA (Figure 2(b), right panel). We consider that the formation of band 3 can have two possible explanations. Either it reflects the fact that the AB didomain interacts with hcDNA more weakly than HMGB1 and is simply the indication of a partial destabilization of the complex that occurs at low protein concentration or at a high concentration of competitor DNA; or the complexes of HMGB1 with hcDNA result from an extremely cooperative binding of two protein molecules which does not occur with the didomain AB. While such a cooperative effect cannot be completely ruled out, with full-length HMGB1 we have never been able to detect any band resembling complex 3 of AB didomain, even at extremely low concentration of HMGB1 or in the presence of extremely high amounts of competitor DNA, and therefore we would tend to favor the former hypothesis.

To summarize the above results, while both domains A and B exhibit a high affinity for hcDNA, the affinity of the full-length HMGB1 protein is more than three orders of magnitude higher, the presence of the acidic tail with both domains A and B being absolutely required for the highest specificity.

### **Binding of the HMG-box domain of SRY to hcDNA**

To find out whether other proteins containing the HMG-box domain could also interact with hcDNA, binding experiments were carried out with the HMG-box domain of the human sex determination factor SRY. As shown in Figure 4(a) (left panel), the HMG-box domain of SRY could clearly interact with hcDNA, with a band pattern similar to those obtained with the single domains of HMGB1. As the HMG-box domain of SRY is able to bind to four-way DNA junctions (4WJ)<sup>32</sup> and to specific sites on linear DNA<sup>33</sup>, we compared the affinities of this domain to hcDNA and to these two substrates. Figure 4(a) (right panel) shows a comparison of the formation of complexes with the three labeled substrates in the presence of variable amounts of nonspecific competitor DNA (DNA from bacteriophage lambda), and demonstrates that the affinities of the HMG-box domain of SRY for the three substrates studied are in the following order: hcDNA > four-way junctions >



linear fragment. Therefore all HMG-box domains tested showed a clear binding preference and a high affinity for hcDNA. We have also demonstrated that the ubiquitous prokaryotic protein HU could specifically bind hcDNA ( $K_d \sim < 1$  nM), Figure 4(b). Although HU is a structurally and evolutionarily unrelated protein to HMGB1, both proteins share similar binding properties, such as DNA bending, supercoiling, compaction and binding to bent DNA<sup>34</sup>. However, despite their similar binding properties towards most other distorted DNA substrates, neither the single HMG-box domains, nor HU protein, nor p53 (our previous work<sup>27</sup>) could bind hcDNA as specifically as the full-length HMGB1 protein, which prompted us to further study the characteristics of HMGB1-hcDNA interactions and to try to define more precisely the target of HMGB1 on this structure.

### **Binding of HMGB1 to hcDNA is not significantly affected by magnesium**

It has been demonstrated that the high affinity of HMGB1 box A for the four-way junction, which was initially observed with the open form of the junction in the presence of EDTA<sup>35</sup>, was strongly decreased in the presence of magnesium ions when the junction adopts the closed form<sup>36</sup>. Given the magnesium concentration in the cell, this led to questioning the physiological importance of the HMG-box domain interactions with four-way junctions. We therefore asked whether the presence of magnesium could interfere with binding of HMGB1 to hcDNA. Complexes formed in the presence of magnesium were analyzed by electrophoresis on polyacrylamide gels in the presence of 1, 10, 100  $\mu$ M or 1 mM  $Mg^{++}$ . As shown in Figure 5, addition of  $Mg^{++}$  up to 1 mM (the highest concentration studied) did not significantly affect the affinity of HMGB1 to hcDNA. A similar result was obtained with the HMG-box domain of SRY (data not shown). The decrease in the intensity of band C2, (which presumably corresponds to the binding of a second HMGB1 molecule to hcDNA as compared to band C1), was the only visible effect of increasing the magnesium concentration. This decrease in affinity of the second binding site may be due to a change of conformation of hcDNA upon addition of magnesium, as suggested by the marked change of mobility of free hcDNA relative to the linear fragment that is observed here at higher magnesium concentrations.

### **Is it the DNA hemicatenane or the DNA loop that determines the high affinity of HMGB1 for hcDNA?**

Given the presence of a small DNA loop in the hcDNA structure, and given the known affinity of HMGB1 for DNA minicircles, we studied whether binding of HMGB1 to hcDNA was due to the presence of a DNA loop, or to the hemicatenane, or to the combined presence of both features.

To try to address this question, we first assumed that the DNA structure was similar in minicircles and in the loop of hcDNA. DNA minicircles with sizes varying between 66 and 87 bp were prepared either from DNA fragments with non-repetitive sequences or from fragments containing a 62 bp

tract of poly(CA)·poly(TG), and used in binding studies performed in parallel with hcDNA. Figure 6 demonstrates a clear binding preference of HMGB1 for hcDNA as compared to 66 bp minicircles of mixed sequence (Figure 6(a)) and 69 bp CA/TG-containing minicircles (Figure 6(b)). The  $K_d$  observed for HMGB1 binding to minicircles is  $\sim 50$  pM, corresponding to an affinity constant of  $2 \times 10^{10}$  in good agreement with the published values of  $10^9$ - $10^{10}$  (refs.<sup>37,38</sup>). Other minicircles of 83 and 87 bp containing mixed sequences and 86 bp containing a CA/TG tract gave very similar results (data not shown) and did not reveal any sequence preference of HMGB1 for minicircles in the 66-87 bp size range. The binding constant of HMGB1 for minicircles must be compared to the value of  $5 \times 10^{12}$  found previously for binding to hcDNA in experiments using a DNA concentration lower than here, 0.16 pM<sup>24</sup>. This ratio of more than two orders of magnitude between binding constants is a first indication of the importance of the DNA hemicatenane in the high affinity of HMGB1 for hcDNA.

A further indication was obtained by studying a mutant of HMGB1 in which all three intercalating residues were mutated to alanine, F<sub>38</sub>, F<sub>103</sub>, I<sub>122</sub> (see Figure 2 above). Figure 6(c) shows that the mutated protein retained almost completely the affinity of the wild-type protein for hcDNA, only the formation of band C2 being strongly impaired, which presumably corresponds to the binding of a second protein molecule to the complex. In strong contrast, the affinity of the mutated protein to 66 bp minicircles was markedly diminished, only the first complex was formed, with an affinity constant decreasing by two orders of magnitude. The fact that the mutant protein binds strongly to hcDNA while binding only weakly to minicircles indicates that the DNA loop is not the main determinant of HMGB1 specificity for hcDNA. Instead it suggests strongly that the high affinity of HMGB1 is due to the presence of a hemicatenane in the structure.

### **HMGB1 protection of DNA hemicatenane from nuclease digestion**

To further study whether HMGB1 binds to the hemicatenane or to the loop, we also performed footprinting experiments by digesting hcDNA with nucleases in the presence of HMGB1. The results obtained with exonuclease III digestion were particularly clear and striking and are shown in Figure 7. First, the presence of a DNA knot was not a significant obstacle to the progression of the enzyme, as exonuclease III digestion in the absence of HMGB1 was only very weakly influenced by the hemicatenane (under the digestion conditions used no pause of the enzyme was observed on linear DNA, data not shown). In contrast, in the presence of HMGB1 strong pauses of the exonuclease were observed on both sides of the molecule immediately before the border between the terminal non-repetitive sequences and the poly(CA)·poly(TG) tract, i.e. at the location of the hemicatenane. This result strongly suggests that HMGB1 binds at this position and prevents the progression of exo III, either directly by steric hindrance, or indirectly by stabilizing the knot. In

either case, the results do not suggest that HMGB1 binds to the DNA loop but are rather in good agreement with a specific binding of HMGB1 to the hemicatenane.

### **Modification of DNA minicircles conformation upon HMGB1 binding**

The high affinity of HMGB1 to minicircles was originally considered to be due to the strong curvature of DNA in such circles<sup>16,17</sup>. However the DNA structure in minicircles much shorter than the DNA persistence length is not precisely known, and chemical footprinting experiments<sup>38</sup> have shown that sites with a strongly modified structure are present in 75 bp minicircles and might represent the actual determinants of HMGB1 affinity. To try to better understand the affinity of HMGB1 for circular DNA, we studied its interactions with minicircles obtained with DNA fragments longer than the persistence length of DNA. Following circularization of a 258 bp fragment from plasmid pBR322, relaxed and negatively supercoiled topoisomers were individually purified by separation on a polyacrylamide gel in the presence of chloroquine (Figure 8(a), see Materials and Methods for a description of the protocol for chloroquine removal from DNA). The interactions of HMGB1 with the purified topoisomers (Figure 8(b)) showed two unexpected characteristics. First, HMGB1 binding resulted in the unusual acceleration of the bands, not in a retardation as generally observed for protein binding in band shift assays. The presence of HMGB1 in the accelerated bands was confirmed by supershifting with anti-HMGB1 antibodies (data not shown), and by the fact that addition of 0.1% SDS to the samples immediately before loading suppressed the accelerated bands (not shown). The second characteristic of these interactions was that all topoisomers did not interact similarly with HMGB1, the relaxed ones forming no complex, the most supercoiled ones binding HMGB1 only weakly, while the highest affinity was observed with topoisomers with linking number differences of  $-2$  and  $-3$  relative to the relaxed minicircles. This unexpected acceleration of the bands, which can only be explained by a change of conformation of the minicircles being both induced by and required for HMGB1 binding, will be further discussed below (see Discussion).

### **Absence of binding of HMGB1 to Z-DNA**

As hcDNA is prepared from DNA fragments containing a CA/TG microsatellite, we performed a similar experiment with a 258 bp DNA fragment containing a 62 bp tract of poly(CA)·poly(TG). Since this sequence is known to be able to adopt the Z-form DNA conformation when subject to topological constraints<sup>39,40</sup>, this also allowed us to test the possibility that Z-DNA was involved in the recognition of hcDNA by HMGB1, as previous reports had suggested that HMGB1 could bind preferentially to brominated poly(CG)·poly(CG) in the Z-DNA conformation<sup>41,42</sup>. The results (Figure 8(c)) with 258 bp minicircles containing CA/TG are even more striking than above: topoisomer  $-1$  was not bound, while all of topoisomer  $-2$  was strongly accelerated, the other topoisomers being all

accelerated upon binding but only weakly bound. As expected, antibodies against Z-DNA (a kind gift of Alexander Rich) bound all the most supercoiled topoisomers (topoisomers -3 to -6), showing the induction of Z-form by supercoiling, but no preferential interaction of HMGB1 with the topoisomers containing Z-DNA could be detected. In addition, using the same assay (data not shown) we did not detect any binding of anti-Z-DNA antibodies to hcDNA.

## Discussion

While HMGB1 has long been known to bind specifically to many distorted DNA structures, the present work shows several unique characteristics of its interactions with hemicatenated DNA loops. First, the affinity of the full-length protein is approximately three orders of magnitude higher than the affinity of its single HMG-box domains or of the HMG-box domain of the sex-determination factor SRY, whereas with most other DNA structures the difference in affinity is much lower. Second, unlike what is usually observed, the AB didomain has a lower affinity for hcDNA than the full-length protein. Third, mutations of the intercalating residues, i.e. of the residues that have been shown to be important for specific binding of distorted DNA structures, have little effect on binding to hcDNA. Fourth, the presence of magnesium ions in the buffer has no significant effect on the interactions, unlike what is observed with four-way junctions. Fifth and most important, HMGB1 affinity for hcDNA is much higher than for the other substrates, the binding constant of  $5 \times 10^{12}$  being of the same order of magnitude as the affinity of highly-specific DNA binding proteins for their recognition sites, such as lac repressor binding to the lac operator or *EcoRI* restriction enzyme binding to its cutting site.

An immediate question raised by these observations is the structural feature of hcDNA recognized by HMGB1. The extremely high affinity of HMGB1 could be due to a strong affinity for DNA hemicatenanes but it could also be due to the distortion of DNA that results from the presence of a very small DNA loop in the structure, or to some other features resulting from the combined presence of both elements, all these possibilities being non-exclusive. Given the fact that hcDNA is derived from fragments containing the poly(CA)·poly(TG) sequence, the possibility was considered that Z-form DNA was present in the structure and involved in this affinity. This possibility can now be ruled out both on the basis of our results and of previous reports in the literature. While the suggestion of a preference of HMGB1 and HMGB2 for Z-DNA had been made previously<sup>41,42</sup>, several articles showed that this was rather a preference of the proteins for brominated poly(CG)·poly(CG) than a general preference for the Z-DNA conformation by itself<sup>42,44</sup>. The fact that we were not able to detect a strong interaction of HMGB1 with Z-DNA in supercoiled DNA circles containing the CA/TG sequence is in good agreement with the latter reports.

HMGB1 and other HMG-box domains have been shown to bind DNA circles of 75-100 bp with affinity constants in the order of  $10^9$ - $10^{10}$  (refs.<sup>37,38</sup>). Since hcDNA was prepared from DNA fragments containing a 62 bp tract of CA/TG, the DNA loop that it contains might be one of the determinants of HMGB1 affinity. On the other hand, the structural feature of DNA minicircles that is recognized by HMGB1 is not precisely known. It was initially assumed that minicircles had a regular circular structure and that HMGB1 binding was simply due to DNA bending, but the results

obtained recently using osmium tetroxide and diethyl pyrocarbonate<sup>38</sup> suggest that HMGB1 binds to alternative DNA structures that occur in minicircles much shorter than the persistence length of DNA. Similarly, the pattern that we observe for HMGB1 binding to topoisomers of DNA minicircles longer than the DNA persistence length does not suggest a mere binding to curved DNA. The fact that accelerated bands appear upon HMGB1 binding to the moderately supercoiled topoisomers, but not to the most supercoiled ones, points to a major change of conformation of the circles, resulting in a more compact form that migrates faster despite the presence of bound HMGB1. A similar observation has been done previously for binding of the archaeobacterial histone-like protein MC1 to 207 bp minicircles<sup>45</sup>. It is interesting to note that the preferential binding to -2 and -3 topoisomers would correspond to a DNA helical twist of  $\sim 11.5$  bp/turn, which is in accord with the change in helical twist observed for DNA binding by the *Drosophila* homolog of HMGB1, HMGD<sup>46,47</sup>. By analogy with the fast migration of the supercoiled forms of larger circular DNA molecules such as plasmids, it could also be suggested that HMGB1 binding induces and/or stabilizes the writhing of the minicircles, possibly by formation of an 8-shaped conformation. In this hypothesis the binding site of HMGB1 would be at the DNA crossover, i.e. at a structure similar to four-way junctions and hemicatenanes.

The mutated form of HMGB1 in which all three intercalating residues have been replaced by alanines retains a strong binding affinity to hcDNA but has only a weak affinity to minicircles. This is a strong indication that the affinity of the triple mutant for hcDNA (and therefore also the affinity of wild-type HMGB1) is not due to the presence of the loop but to the hemicatenane. This is strengthened by the two orders of magnitude difference between binding affinity constants of HMGB1 to hcDNA and to minicircles irrespective of the presence or absence of the CA/TG sequence, as well as by the exonuclease III footprinting experiments revealing a protection of the hemicatenane site by HMGB1. Further proof that the hemicatenane is the determinant of HMGB1 binding to hcDNA will require the preparation of hemicatenanes of two independent DNA fragments, no longer connected by a loop. We have prepared hcDNA by using a DNA fragment with a short restriction site inserted within the repetitive sequence, in order to cut hcDNA in the loop and to obtain true hemicatenanes. We found that the presence of such a site did not interfere with hcDNA formation. However we have observed, first, that such restriction sites become difficult to cut when present in the loop, and, second, that once the loop is cut the hemicatenane becomes extremely unstable even in the presence of HMGB1 (C.G., L.d.K., and F.S., unpublished results) and readily dissociates into two duplexes. Therefore other ways will have to be designed in order to prepare hemicatenanes of distinct DNA molecules, perhaps using the procedure suggested recently by Bucka and Stasiak for making knotted DNA circles<sup>48</sup>.

Since the DNA hemicatenane and the four-way junction both consist in the junction of two double-stranded DNA molecules, they could be expected to share some degree of structural similarity. In particular, our finding that the DNA-intercalating residue of the HMGB1 domain A (Phe38), rather than those of the domain B, is involved in high-affinity binding to hcDNA is in accord with previous reports on the dominant role of domain A in the *structure*-specific binding of HMGB1 protein to four-way-junctions, DNA minicircles, or bent (cisplatin-modified) DNA<sup>18,37,49,50,51,52,53,54</sup>. Interestingly, the presence of the C-terminal flanking sequence within the domain B was sufficient for a strong binding of the domain to hcDNA, with no further increase in binding upon attachment of the extended N-terminus (the <sup>85</sup>TKKKFKD<sup>91</sup> sequence), unlike what had been reported for DNA bending or binding of the domain B to bent (cisplatin-modified) DNA<sup>20,54</sup>.

Similarly to results obtained with SRY by other authors, data shown here were obtained with a domain that was only defined on the basis of sequence comparisons, the influence of flanking sequences having not been investigated. With HMGB1 in contrast, the marked influence of amino-acid residues adjacent to the HMG-box domains has been well documented, for example the influence of the seven amino-acids on the N-terminal side of domain B on DNA binding and bending is extremely striking<sup>20,54</sup>. A similar investigation would thus deserve to be done not only with SRY but also with other members of the HMG-box protein family. In particular some proteins of the family, such as UBF, contain several HMG-box domains. If the situation is similar to HMGB1 where the presence of the two domains is required for maximum affinity, then results with a combination of domains or with the full-length proteins might be significantly different from results obtained with single domains.

It has been suggested that the affinity of the HMG-box for four-way junctions was the result of a structural coincidence with yet unknown biological binding sites, and the functional significance of these interactions was questioned based on the strong decrease in binding specificity observed in the presence of magnesium, i.e. under more relevant physiological conditions<sup>36</sup>. In contrast, HMGB1 and HMG-box domains consistently show a high affinity for binding to hcDNA even in the presence of high concentrations of Mg<sup>++</sup>, conditions under which binding to four-way junctions is no longer detectable. Our results thus point to an exceptional structural fit between HMGB1 and hemicatenanes, even under physiological conditions, which suggests that the question as to whether hemicatenanes exist in the cell should be thoroughly investigated, a suggestion strengthened by our recent observation of a high affinity of tumor suppressor p53 for hcDNA<sup>27</sup>. While there is no doubt that the majority of DNA in the genome is in a double-stranded linear structure of the B-DNA type, the possibility clearly exists for unfrequent alternative DNA structures to be present. On one hand, hemicatenanes and more generally DNA knots are of great interest as potential intermediates during

processes such as recombination or replication that involve more than two strands of DNA. In particular a role for hemicatenanes in genome function has previously been suggested in relation with recombination<sup>55-60</sup> or replication<sup>61-65</sup>, and several studies using two-dimensional electrophoresis on agarose gels under native conditions have shown the existence of junctions between DNA molecules *in vivo* that could be either Holliday junctions or hemicatenanes<sup>64-69</sup>. In addition, a very attractive possibility would be that hemicatenanes and HMGB1 participate in the general organization of the genome in chromosomal loops and domains, a suggestion strengthened by the recent and striking observation of a role of HMGB1 and its yeast homolog NHP6A/B in the maintenance of genome stability<sup>70</sup>.



## Materials and Methods

### DNA

Hemicatenated DNA loops were prepared with a 120 bp *EcoRI-ClaI* fragment from plasmid pE10 that contains a 62 bp tract of poly(CA)·poly(TG) inserted in plasmid pUC19 (accession number [X96980](#), see fragment map Figure 1). The linear DNA fragment containing SRY binding sites was the 279 bp *Sau96I* fragment from plasmid pTZ19R, which contains four copies of the consensus recognition sequence of SRY<sup>33</sup>. 258 bp DNA minicircles containing no repetitive sequence were prepared using the 258 bp *Sau3AI* fragment of plasmid pBR322. 258 bp DNA minicircles containing the poly(CA)·poly(TG) sequence were obtained using the 258 bp *HinP1I-ClaI* fragment of plasmid pE10. DNA fragments were purified by preparative polyacrylamide gel electrophoresis and electroelution, dephosphorylated with calf intestine phosphatase, 5'-end labeled with polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]-ATP. Topoisomers of 258 bp DNA minicircles were obtained by classical techniques involving circularization of linear fragments with T4 DNA ligase in the presence of variable amounts of ethidium bromide, followed by ethidium bromide removal and determination of DNA linking numbers by electrophoresis on polyacrylamide gels in the presence of 20  $\mu$ g/ml chloroquine<sup>71,72</sup>. 66 bp minicircles of non-repeated sequence were prepared by ligation of 66-bp duplexes with sticky ends in the presence of protein HMGB1 as detailed<sup>20,27</sup>. 69 bp minicircles containing a 62 bp tract of poly(CA)·poly(TG) were prepared similarly by circularization of the 69 bp *Sau3AI* fragment from plasmid pE10. Four-way DNA junctions were prepared using synthetic oligonucleotides with sequences identical to the nucleotides used by Bianchi *et al.*<sup>35</sup>, one of the four oligonucleotides being 5' end-labeled with <sup>32</sup>P. After association of the four strands, four-way junctions were purified on a polyacrylamide gel and electroeluted.

### Hemicatenated DNA loops

The procedure for hcDNA preparation was described previously<sup>23</sup>. In brief, a 120 bp DNA fragment containing a 62 bp tract of poly(CA)·poly(TG) flanked by non-repetitive sequences is heat-denatured and the strands are allowed to reassociate at 37° in 50 mM NaCl, 25 mM Tris-HCl pH 7.5, 1 mM EDTA, in the presence of protein HMGB1 at a concentration of  $\sim$ 5 ng/ $\mu$ l. The complexes of hcDNA with HMGB1 are then purified by electrophoresis on a polyacrylamide gel in native conditions and electroeluted. After protein removal by chloroform extraction in the presence of 1% SDS and 1M NaCl, hcDNA is recovered by ethanol precipitation in the presence of polyacrylamide carrier<sup>73</sup>, dissolved in 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and stored at 4°C in low-binding polypropylene tubes (Axygen)<sup>74</sup>.

### Purification of DNA topoisomers

Topoisomers, obtained by circularization of labeled DNA fragments in the presence of ethidium bromide, were individually purified by preparative electrophoresis on 4% polyacrylamide gels (in 6.7 mM Tris, 3.3 mM Na-acetate, 1 mM EDTA, pH 7.8, run at 4°C with buffer recirculation) containing chloroquine at a concentration of 20  $\mu\text{g/ml}$  both in the gel and in the buffer. After autoradiography, bands were cut and topoisomers were electroeluted and ethanol precipitated. As chloroquine is insoluble in ethanol, it coprecipitated with DNA and interfered strongly with further experiments (DNA-protein interactions, restriction enzyme digestion, etc). Since chloroquine does not bind DNA very strongly ( $K_d > 25 \mu\text{M}$ <sup>75</sup>), it could be removed from DNA by a second preparative electrophoresis, this time on a polyacrylamide gel containing no chloroquine (4% polyacrylamide gel in Tris-acetate-EDTA buffer, at room temperature, without buffer recirculation). After electroelution and ethanol precipitation, the absence of chloroquine from the purified topoisomers was confirmed by their ability to be digested with restriction enzymes (data not shown), to interact with proteins, and by their normal migration on gels containing no chloroquine (see Figure 8).

## Proteins

Non-recombinant HMGB1 protein was purified from cultured monkey cells (CV1 line) or from HeLa cells as described<sup>76</sup>.

HMGB1 domain A (residues 1-87), domain B (residues 85-180), didomain AB (residues 1-180), full-length protein, and mutants of DNA-intercalating residues of domains A and B were produced in *Escherichia coli* as (His)<sub>6</sub>-fusion proteins, after cloning the corresponding sequences into vector pET15b (Novagen) which contains a promoter for T7 RNA polymerase and sequences required for expression of fusion proteins having a polyhistidine tag at their N-terminus. Alanine mutagenesis of residues Phe38 (domain A) and Phe103, I122 or Phe103+I122 (domain B) was carried out by the protocol using “chimeras” of the HMG-box domains (M.Š., paper in preparation). After verification of the constructs by DNA sequencing, the (His)<sub>6</sub>-fusion proteins were produced in *E. coli* strain BL21(DE3)pLysS, extracted under non-denaturing conditions and affinity purified on Ni-NTA spin columns as recommended in the Qiagen instruction manual, and further purified by FPLC chromatography on a Mono-S column. Alternatively, the AB didomain was prepared by direct cleavage of HMGB1 from calf thymus with trypsin and purification on a Mono-S column<sup>18</sup>.

The HMG-box domain of SRY was similarly expressed in *E. coli* as a (His)<sub>6</sub>-fusion protein by cloning the corresponding DNA sequence from plasmid pGEX-2T-SRY containing the human sry gene (a gift of Marc Fellous and Laura Salas-Cortes) into vector pET15b.

Protein HU was the kind gift of Josette Rouvière-Yaniv.

## **Determination of DNA and protein concentrations.**

The estimation of binding constants required measurements of DNA and protein concentrations as accurate as possible. DNA fragments were gel-purified in amounts larger than 10 $\mu$ g and their concentration was measured spectrophotometrically before kinase labelling. After radioactive labelling of known amounts of fragments, followed by removal of all unincorporated radioactivity by chloroform extraction and quantitative ethanol precipitation in the presence of linear polyacrylamide carrier<sup>73</sup>, DNA samples were counted in a scintillation counter and specific activities in cpm/ $\mu$ g calculated. In such a way, DNA amounts could always be known with good precision by simple counting of the samples. Protein concentrations were determined spectrophotometrically. Given the uncertainties on molar extinction coefficients in the case of recombinant and mutant proteins, protein concentrations were confirmed by electrophoresis of variable amounts of the purified proteins on SDS-polyacrylamide gels, followed by staining with Coomassie blue, gel scanning with a Molecular Dynamics densitometer, and comparison of the data with a calibration curve obtained by applying the same procedure on the same gel to serial dilutions of a known amount of bovine serum albumin or of carbonic anhydrase.

## **DNA-protein interactions**

DNA-protein complexes were formed using a known concentration of labelled DNA as indicated in the Figure legends (usually corresponding to ~1000-5000 cpm of <sup>32</sup>P-labelled DNA per sample), plus DNA from *E. coli* or from bacteriophage lambda used as a non-radioactive competitor when indicated. Incubations were in 25  $\mu$ l of 50 mM NaCl, 25 mM Tris-HCl pH 7.5, 1 mM DTT, 1 mM EDTA, plus 0.1% Triton X-100 and bovine serum albumin at 100  $\mu$ g/ml to prevent any protein adsorption on tube walls. Samples were incubated at 37° for 30 min and loaded on a pre-electrophoresed polyacrylamide gel without addition of tracking dye. For experiments in the presence of magnesium, the EDTA concentration in the samples was kept below 0.1 mM and MgCl<sub>2</sub> was added to 1 mM to the incubation buffer. For exonuclease III digestion, complexes formed as described above were incubated with 10 units of enzyme for 5 min at 25° in the presence of 2 mM MgCl<sub>2</sub>.

## **Gel electrophoresis**

DNA-protein complexes were analyzed by electrophoresis on 4% polyacrylamide gels (acrylamide:bis-acrylamide 29:1) in 6.7 mM Tris, 3.3 mM Na acetate, 1 mM EDTA, pH 7.8, at 4° with buffer recirculation. Pre-electrophoresis was for 2 hr and migration of the samples was for 4 hr at 10 V/cm. Gels in the presence of magnesium were identical except that EDTA was not used and MgCl<sub>2</sub> at the indicated concentration was present in the gel and in the buffer.

## **Acknowledgements**

We thank Alexander Rich for the antibodies against Z-DNA, Josette Rouvière-Yaniv for protein HU, Marco Bianchi for the plasmid containing the rat HMGB1 gene, and Laura Salas-Cortes and Marc Fellous for the plasmid containing the human *sry* gene. We also thank Caroline Perrin for technical assistance and Alena Bačíková for purification of HMGB1 protein, domains, and mutants.

This research was supported by the Centre National de la Recherche Scientifique (CNRS), Université Pierre et Marie Curie (Paris 6), Université Denis Diderot (Paris 7), by the international collaboration program between CNRS and the Academy of Sciences of the Czech Republic, and by grants to M.Š. from the Academy of Sciences of the Czech Republic (AVOZ50040507: Biophysics of dynamic structures of biological systems) and the Grant Agency of the Czech Republic (204/05/2031).

## References

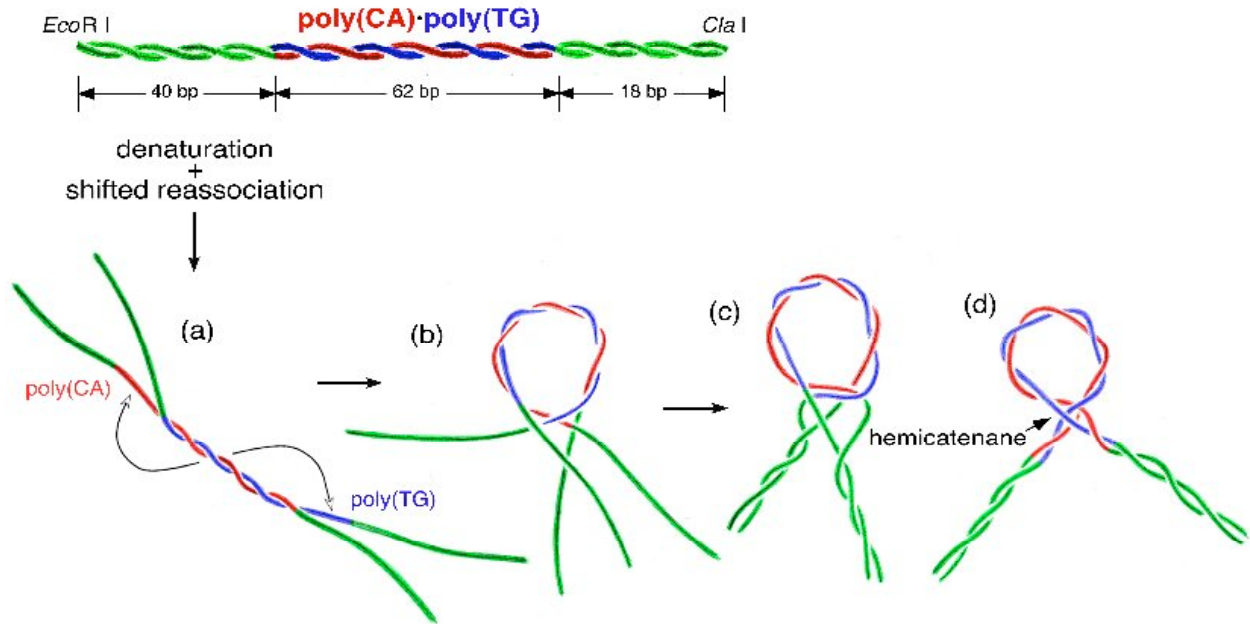
1. Bustin, M. (2001). Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem. Sci.* **26**, 152-153.
2. Baxevasis, A. D. & Landsman, D. (1995). The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res.* **23**, 1604-1613.
3. Wegner, M. (1999). From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res.* **27**, 1409-1420.
4. Bianchi, M. E. & Beltrame, M. (1998). Flexing DNA: HMG-box proteins and their partners. *Am. J. Hum. Genet.* **63**, 1573-1577.
5. Agresti, A. & Bianchi, M. E. (2003). HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* **13**, 170-178.
6. Travers, A. (2000). Recognition of distorted DNA structures by HMG domains. *Curr. Opin. Struct. Biol.* **10**, 102-109.
7. Thomas, J. O. & Travers, A. A. (2001). HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem. Sci.* **26**, 167-174.
8. Thomas, J. O. (2001). HMG1 and 2: architectural DNA-binding proteins. *Biochem. Soc. Trans.* **29**, 395-401.
9. Bustin, M. (1999). Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol. Cell. Biol.* **19**, 5237-5246.
10. Bonaldi, T., Langst, G., Strohner, R., Becker, P. B. & Bianchi, M. E. (2002). The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. *EMBO J.* **21**, 6865-6873.
11. Travers, A. A. (2003). Priming the nucleosome: a role for HMGB proteins? *EMBO Rep.* **4**, 131-136.
12. Wang, H., Bloom, O., Zhang, M., Vishnubhakat, J. M., Ombrellino, M., Che, J., Frazier, A., Yang, H., Ivanova, S., Borovikova, L., Manogue, K. R., Faist, E., Abraham, E., Andersson, J., Andersson, U., Molina, P. E., Abumrad, N. N., Sama, A. & Tracey, K. J. (1999). HMG-1 as a late mediator of endotoxin lethality in mice. *Science* **285**, 248-251.
13. Taguchi, A., Blood, D. C., del Toro, G., Canet, A., Lee, D. C., Qu, W., Tanji, N., Lu, Y., Lalla, E., Fu, C., Hofmann, M. A., Kislinger, T., Ingram, M., Lu, A., Tanaka, H., Hori, O., Ogawa, S., Stern, D. M. & Schmidt, A. M. (2000). Blockade of RAGE-amphoterin signalling suppresses tumour growth and metastases. *Nature* **405**, 354-360.
14. Lotze, M. T. & Tracey, K. J. (2005). High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nat Rev Immunol* **5**, 331-342.
15. Grasser, K. D. (2003). Chromatin-associated HMGA and HMGB proteins: versatile co-regulators of DNA-dependent processes. *Plant Mol. Biol.* **53**, 281-295.
16. Paull, T. T., Haykinson, M. J. & Johnson, R. C. (1993). The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. *Genes Dev.* **7**, 1521-1534.
17. Pil, P. M., Chow, C. S. & Lippard, S. J. (1993). High-mobility-group 1 protein mediates DNA bending as determined by ring closures. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 9465-9469.
18. Štros, M., Štokrová, J. & Thomas, J. O. (1994). DNA looping by the HMG-box domains of HMG1 and modulation of DNA binding by the acidic C-terminal domain. *Nucleic Acids Res.* **22**, 1044-1051.
19. Teo, S. H., Grasser, K. D. & Thomas, J. O. (1995). Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *Eur. J. Biochem.* **230**, 943-950.
20. Štros, M. (1998). DNA bending by the chromosomal protein HMG1 and its high mobility group box domains. Effect of flanking sequences. *J. Biol. Chem.* **273**, 10355-10361.
21. Sheflin, L. G. & Spaulding, S. W. (1989). High mobility group protein 1 preferentially conserves torsion in negatively supercoiled DNA. *Biochemistry* **28**, 5658-5664.

22. Sheflin, L. G., Fucile, N. W. & Spaulding, S. W. (1993). The specific interactions of HMG 1 and 2 with negatively supercoiled DNA are modulated by their acidic C-terminal domains and involve cysteine residues in their HMG 1/2 boxes. *Biochemistry* **32**, 3238-3248.
23. Gaillard, C. & Strauss, F. (2000). DNA loops and semicatenated DNA junctions. *BMC Biochem.* **1**, 1.
24. Gaillard, C. & Strauss, F. (2000). High affinity binding of proteins HMG1 and HMG2 to semicatenated DNA loops. *BMC Mol. Biol.* **1**, 1.
25. Lyubchenko, Y. L., Shlyakhtenko, L. S., Binus, M., Gaillard, C. & Strauss, F. (2002). Visualization of hemiknot DNA structure with an atomic force microscope. *Nucleic Acids Res.* **30**, 4902-4909.
26. Gaillard, C., Shlyakhtenko, L. S., Lyubchenko, Y. L. & Strauss, F. (2002). Structural analysis of hemicatenated DNA loops. *BMC Struct. Biol.* **2**, 7.
27. Štros, M., Muselíková-Polanská, E., Pospíšilová, S. & Strauss, F. (2004). High-affinity binding of tumor-suppressor protein p53 and HMGB1 to hemicatenated DNA loops. *Biochemistry* **43**, 7215-7225.
28. Wong, B. C., Chiu, S. K. & Chow, S. A. (1998). The role of negative superhelicity and length of homology in the formation of paranemic joints promoted by RecA protein. *J. Biol. Chem.* **273**, 12120-12127.
29. He, Q., Ohndorf, U. M. & Lippard, S. J. (2000). Intercalating residues determine the mode of HMG1 domains A and B binding to cisplatin-modified DNA. *Biochemistry* **39**, 14426-14435.
30. Štros, M. & Muselíková, E. (2000). A role of basic residues and the putative intercalating phenylalanine of the HMG-1 box B in DNA supercoiling and binding to four-way DNA junctions. *J. Biol. Chem.* **275**, 35699-35707.
31. Yen, Y. M., Wong, B. & Johnson, R. C. (1998). Determinants of DNA binding and bending by the *Saccharomyces cerevisiae* high mobility group protein NHP6A that are important for its biological activities. Role of the unique N terminus and putative intercalating methionine. *J. Biol. Chem.* **273**, 4424-4435.
32. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R. & Bianchi, M. E. (1992). SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J.* **11**, 4497-4506.
33. Harley, V. R., Lovell-Badge, R. & Goodfellow, P. N. (1994). Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res.* **22**, 1500-1501.
34. Swinger, K. K. & Rice, P. A. (2004). IHF and HU: flexible architects of bent DNA. *Curr. Opin. Struct. Biol.* **14**, 28-35.
35. Bianchi, M. E., Beltrame, M. & Paonessa, G. (1989). Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **243**, 1056-1059.
36. Pöhler, J. R., Norman, D. G., Bramham, J., Bianchi, M. E. & Lilley, D. M. (1998). HMG box proteins bind to four-way DNA junctions in their open conformation. *EMBO J.* **17**, 817-826.
37. Webb, M., Payet, D., Lee, K. B., Travers, A. A. & Thomas, J. O. (2001). Structural requirements for cooperative binding of HMG1 to DNA minicircles. *J. Mol. Biol.* **309**, 79-88.
38. Nakamura, Y., Shimizu, M. & Yoshida, M. (2002). Distorted DNA structures induced by HMGB2 possess a high affinity for HMGB2. *J. Biochem. (Tokyo)*. **131**, 153-160.
39. Haniford, D. B. & Pulleyblank, D. E. (1983). Facile transition of poly[d(TG) x d(CA)] into a left-handed helix in physiological conditions. *Nature* **302**, 632-634.
40. Nordheim, A. & Rich, A. (1983). The sequence (dC-dA)<sub>n</sub> X (dG-dT)<sub>n</sub> forms left-handed Z-DNA in negatively supercoiled plasmids. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 1821-1825.
41. Gut, S. H., Bischoff, M., Hobi, R. & Kuenzle, C. C. (1987). Z-DNA-binding proteins from bull testis. *Nucleic Acids Res.* **15**, 9691-9705.
42. Christen, T., Bischoff, M., Hobi, R. & Kuenzle, C. C. (1990). High mobility group proteins 1 and 2 bind preferentially to brominated poly(dG-dC).poly(dG-dC) in the Z-DNA conformation but not to other types of Z-DNA. *FEBS Lett.* **267**, 139-141.
43. Hamada, H. & Bustin, M. (1985). Hierarchy of binding sites for chromosomal proteins HMG 1 and 2 in supercoiled deoxyribonucleic acid. *Biochemistry* **24**, 1428-1433.

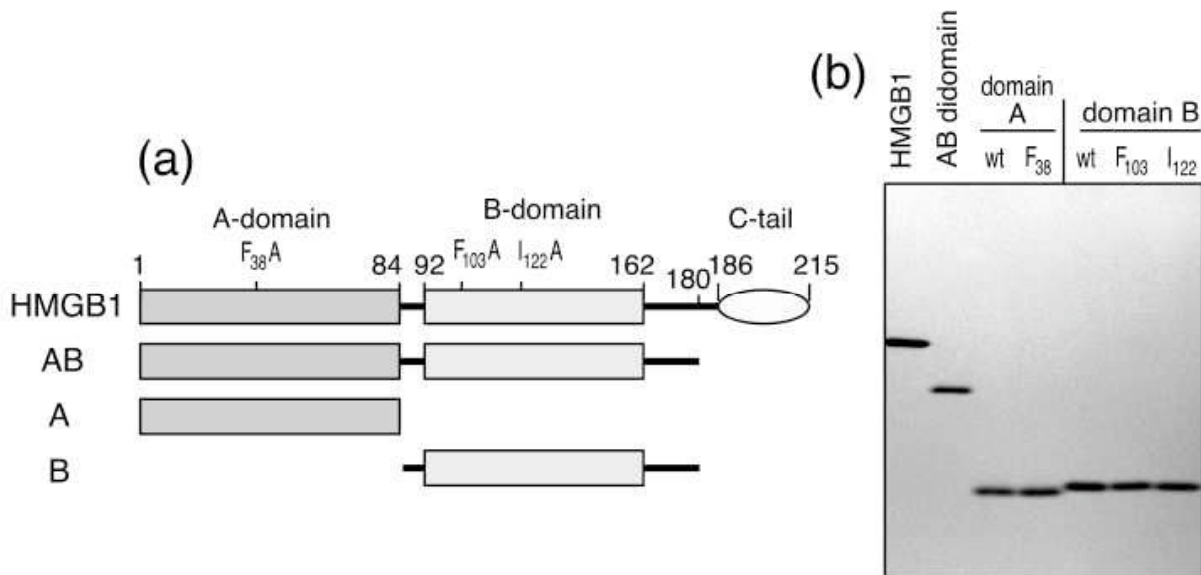
44. Rohner, K. J., Hobi, R. & Kuenzle, C. C. (1990). Z-DNA-binding proteins. Identification critically depends on the proper choice of ligands. *J. Biol. Chem.* **265**, 19112-19115.
45. Teyssier, C., Toulme, F., Touzel, J. P., Gervais, A., Maurizot, J. C. & Culard, F. (1996). Preferential binding of the archaeobacterial histone-like MC1 protein to negatively supercoiled DNA minicircles. *Biochemistry* **35**, 7954-7958.
46. Travers, A. (2000). Recognition of distorted DNA structures by HMG domains. *Curr. Opin. Struct. Biol.* **10**, 102-109.
47. Murphy, F. V. t., Sweet, R. M. & Churchill, M. E. (1999). The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition. *EMBO J.* **18**, 6610-6618.
48. Bucka, A. & Stasiak, A. (2002). Construction and electrophoretic migration of single-stranded DNA knots and catenanes. *Nucleic Acids Res.* **30**, e24.
49. Chow, C. S., Barnes, C. M. & Lippard, S. J. (1995). A single HMG domain in high-mobility group 1 protein binds to DNAs as small as 20 base pairs containing the major cisplatin adduct. *Biochemistry* **34**, 2956-2964.
50. Webb, M. & Thomas, J. O. (1999). Structure-specific binding of the two tandem HMG boxes of HMG1 to four-way junction DNA is mediated by the A domain. *J. Mol. Biol.* **294**, 373-387.
51. Štros, M. & Reich, J. (1998). Formation of large nucleoprotein complexes upon binding of the high-mobility-group (HMG) box B-domain of HMG1 protein to supercoiled DNA. *Eur. J. Biochem.* **251**, 427-434.
52. Teo, S. H., Grasser, K. D., Hardman, C. H., Broadhurst, R. W., Laue, E. D. & Thomas, J. O. (1995). Two mutations in the HMG-box with very different structural consequences provide insights into the nature of binding to four-way junction DNA. *EMBO J.* **14**, 3844-3853.
53. Grasser, K. D., Teo, S. H., Lee, K. B., Broadhurst, R. W., Rees, C., Hardman, C. H. & Thomas, J. O. (1998). DNA-binding properties of the tandem HMG boxes of high-mobility-group protein 1 (HMG1). *Eur. J. Biochem.* **253**, 787-795.
54. Štros, M. (2001). Two mutations of basic residues within the N-terminus of HMG-1 B domain with different effects on DNA supercoiling and binding to bent DNA. *Biochemistry* **40**, 4769-4779.
55. Cunningham, R. P., Wu, A. M., Shibata, T., DasGupta, C. & Radding, C. M. (1981). Homologous pairing and topological linkage of DNA molecules by combined action of E. coli RecA protein and topoisomerase I. *Cell* **24**, 213-223.
56. Bianchi, M., DasGupta, C. & Radding, C. M. (1983). Synapsis and the formation of paranemic joints by E. coli RecA protein. *Cell* **34**, 931-939.
57. Kmiec, E. B. & Holloman, W. K. (1986). Homologous pairing of DNA molecules by Ustilago rec1 protein is promoted by sequences of Z-DNA. *Cell* **44**, 545-554.
58. Umlauf, S. W., Cox, M. M. & Inman, R. B. (1990). Triple-helical DNA pairing intermediates formed by recA protein. *J. Biol. Chem.* **265**, 16898-16912.
59. McKee, B. D., Habera, L. & Vrana, J. A. (1992). Evidence that intergenic spacer repeats of Drosophila melanogaster rRNA genes function as X-Y pairing sites in male meiosis, and a general model for achiasmatic pairing. *Genetics* **132**, 529-544.
60. Polanco, C., Gonzalez, A. I. & Dover, G. A. (2000). Patterns of variation in the intergenic spacers of ribosomal DNA in Drosophila melanogaster support a model for genetic exchanges during X-Y pairing. *Genetics* **155**, 1221-1229.
61. Sogo, J. M., Stahl, H., Koller, T. & Knippers, R. (1986). Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. *J. Mol. Biol.* **189**, 189-204.
62. Laurie, B., Katritch, V., Sogo, J., Koller, T., Dubochet, J. & Stasiak, A. (1998). Geometry and physics of catenanes applied to the study of DNA replication. *Biophys. J.* **74**, 2815-2822.
63. Lucas, I. & Hyrien, O. (2000). Hemicatenanes form upon inhibition of DNA replication. *Nucleic Acids Res.* **28**, 2187-2193.

64. Wellinger, R. E., Schar, P. & Sogo, J. M. (2003). Rad52-independent accumulation of joint circular minichromosomes during S phase in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **23**, 6363-6372.
65. Lopes, M., Cotta-Ramusino, C., Liberi, G. & Foiani, M. (2003). Branch Migrating Sister Chromatid Junctions Form at Replication Origins through Rad51/Rad52-Independent Mechanisms. *Mol. Cell* **12**, 1499-1510.
66. Schwacha, A. & Kleckner, N. (1995). Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell* **83**, 783-791.
67. Kajander, O. A., Karhunen, P. J., Holt, I. J. & Jacobs, H. T. (2001). Prominent mitochondrial DNA recombination intermediates in human heart muscle. *EMBO Rep.* **2**, 1007-1012.
68. Benard, M., Maric, C. & Pierron, G. (2001). DNA replication-dependent formation of joint DNA molecules in *Physarum polycephalum*. *Mol. Cell* **7**, 971-980.
69. Segurado, M., Gomez, M. & Antequera, F. (2002). Increased recombination intermediates and homologous integration hot spots at DNA replication origins. *Mol. Cell* **10**, 907-916.
70. Giavara, S., Kosmidou, E., Hande, M. P., Bianchi, M. E., Morgan, A., d'Adda di Fagagna, F. & Jackson, S. P. (2005). Yeast Nhp6A/B and mammalian Hmgb1 facilitate the maintenance of genome stability. *Curr. Biol.* **15**, 68-72.
71. Shure, M. & Vinograd, J. (1976). The number of superhelical turns in native virion SV40 DNA and minicol DNA determined by the band counting method. *Cell* **8**, 215-226.
72. Shure, M., Pulleyblank, D. E. & Vinograd, J. (1977). The problems of eukaryotic and prokaryotic DNA packaging and in vivo conformation posed by superhelix density heterogeneity. *Nucleic Acids Res.* **4**, 1183-1205.
73. Gaillard, C. & Strauss, F. (1990). Ethanol precipitation of DNA with linear polyacrylamide as carrier. *Nucleic Acids Res.* **18**, 378.
74. Gaillard, C. & Strauss, F. (2000). Eliminating DNA losses and denaturation upon storage in plastic microtubes. *International Biotechnology Laboratory* **18**, 24.
75. Kwakye-Berko, F. & Meshnick, S. R. (1989). Binding of chloroquine to DNA. *Mol. Biochem. Parasitol.* **35**, 51-55.
76. Gaillard, C. & Strauss, F. (1994). Association of poly(CA).poly(TG) DNA fragments into four-stranded complexes bound by HMG1 and 2. *Science* **264**, 433-436.

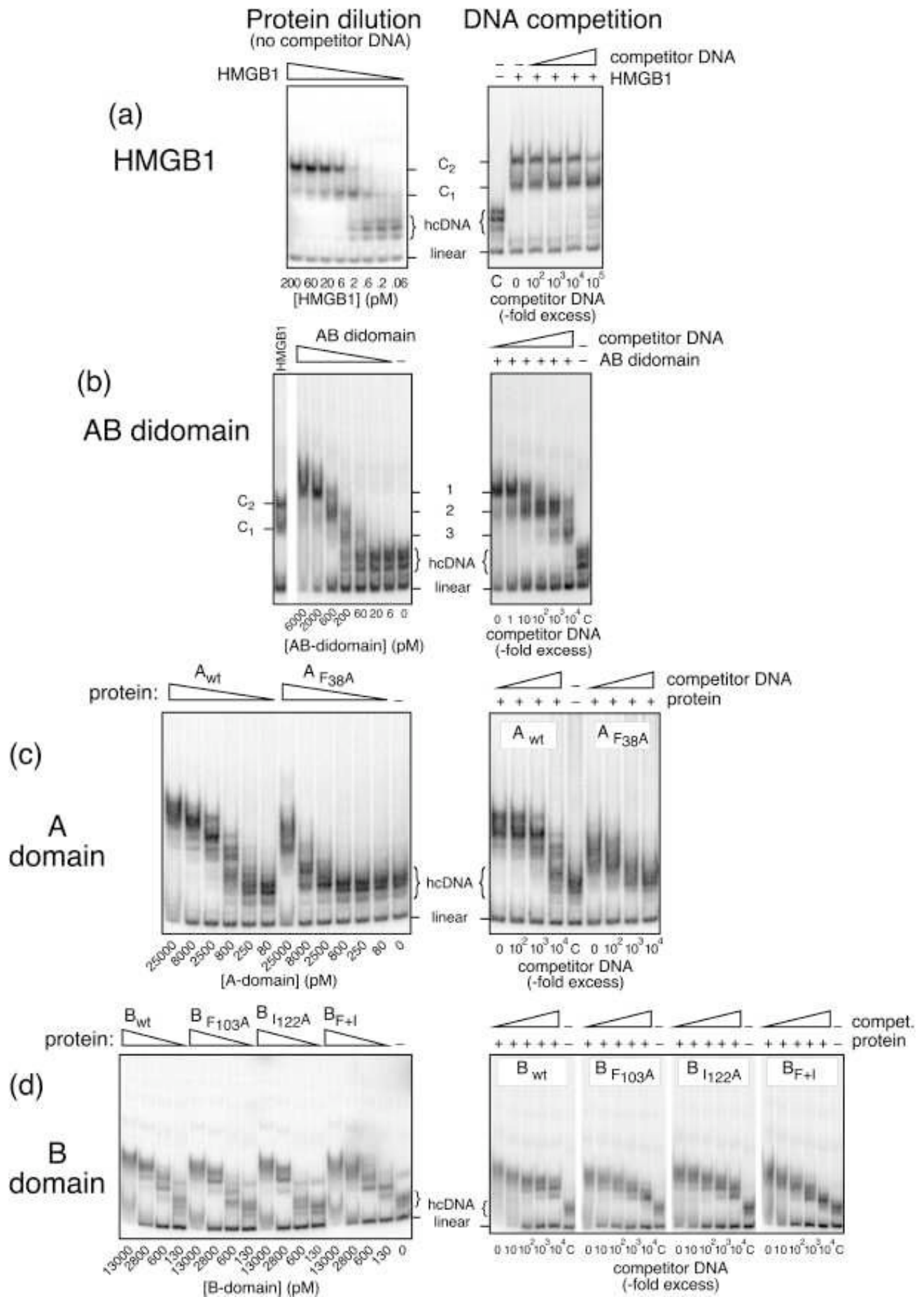




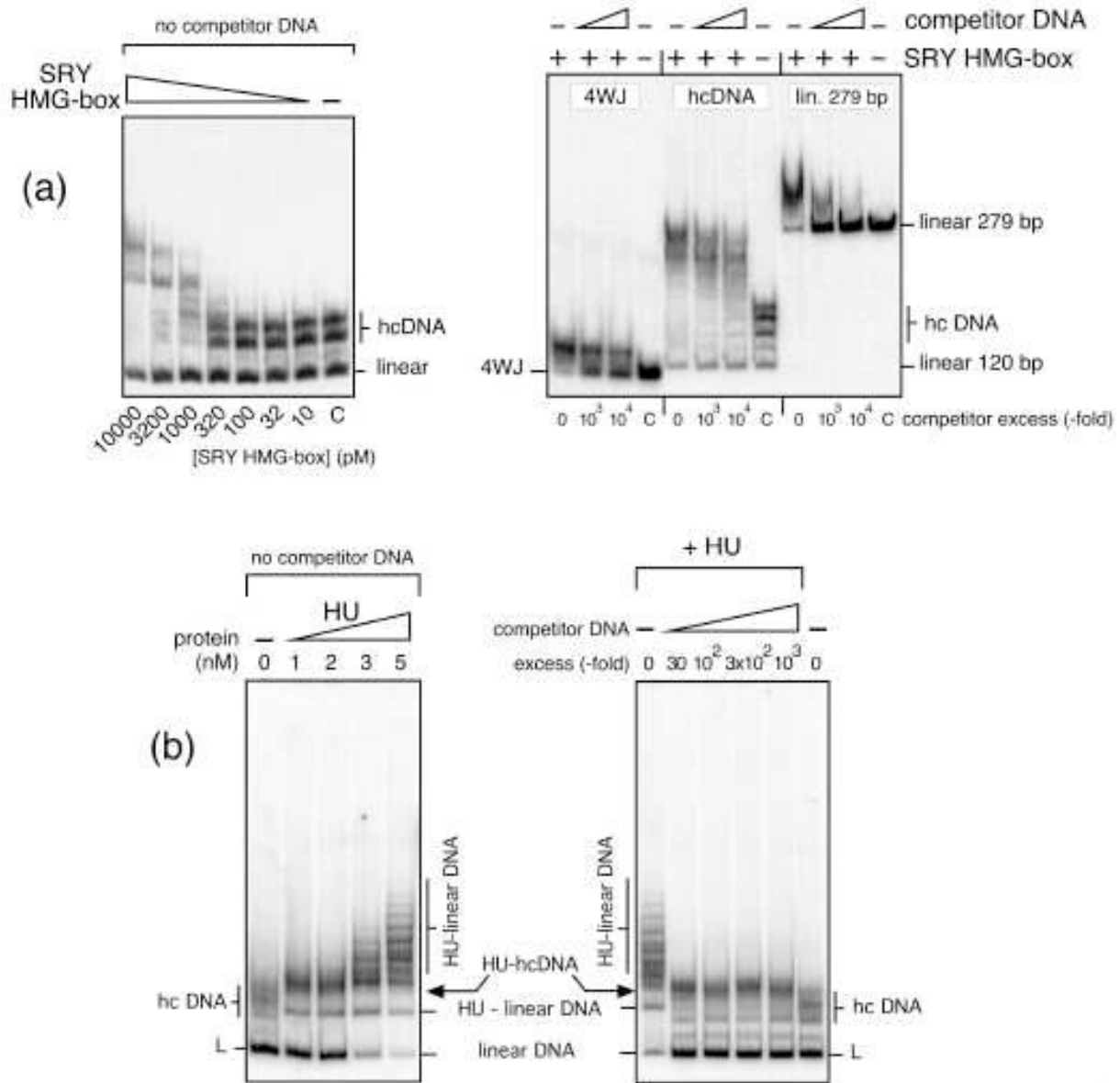
**Figure 1.** Schematic representation of hemicatenated loops and of the model for the mechanism of their formation. A DNA fragment containing a tract of poly(CA)·poly(TG) (red and blue) flanked by non-repetitive sequences (green) is heat-denatured and allowed to reanneal. Pairing of the strands in the repetitive region can occur either in register, reforming the initial double-stranded fragment, or with a shift, forming an intermediate (a) in which poly(CA) and poly(TG) are paired on part of their length only. One of the single-stranded ends can then insert in the fork formed by the two single-strands at the opposite end (b). Pairing of the non-repetitive terminal regions stabilizes the structure (c) and leads to the formation of a hemicatenated loop (d), a structure in which the double helix is folded in a loop maintained at its base by a hemicatenane, i.e. the junction of two DNA duplexes with one strand of one duplex passing between the two strands of the other duplex, and reciprocally.



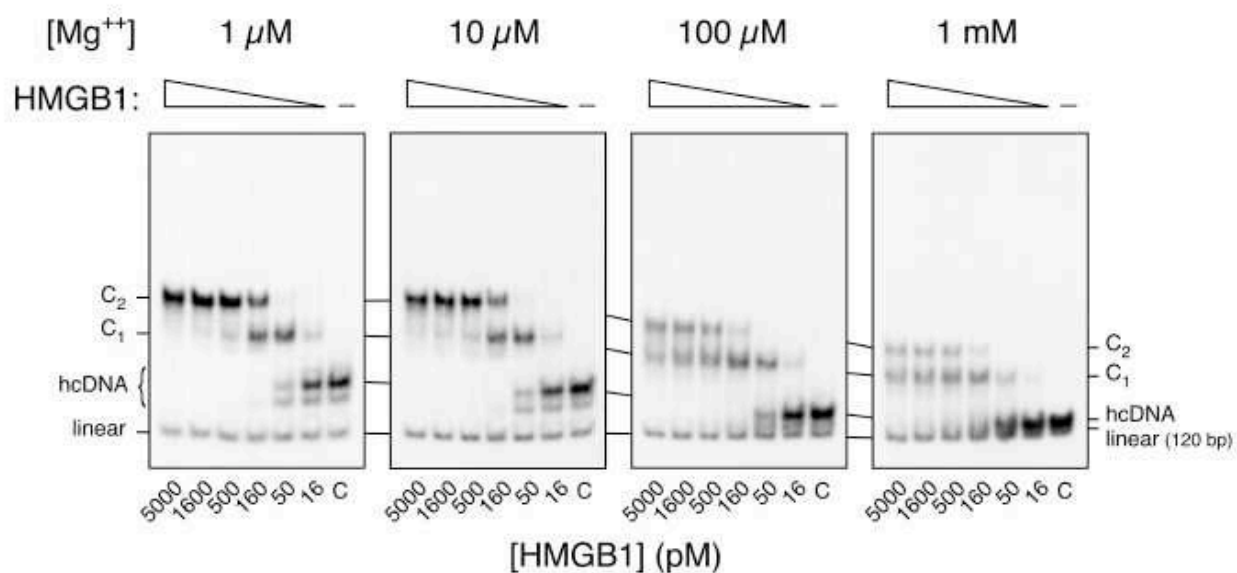
**Figure 2.** Domains of HMGB1 studied in the present work. The domains studied are represented in (a): domain A was strictly limited to the HMG-box sequences, whereas domain B included 7 residues on the N-terminal side and was extended to position 180 on its C-terminal side (originally designated domain B7 in <sup>30,54</sup>). Positions of the amino-acids that were mutated to alanine in the mutants used in this work are indicated, F<sub>38</sub> (box A), and F<sub>103</sub> and I<sub>122</sub> (box B). (b): SDS-polyacrylamide gel of purified proteins, domains and mutants, showing the degree of homogeneity obtained after FPLC purification. Each lane was loaded with about 5  $\mu$ g of highly purified protein, and the gel was stained with coomassie-blue.



**Figure 3.** Interactions of HMGB1, domains, and mutants, with hcDNA; effect of protein concentration (left panels) and of competitor DNA concentration (right panels). The HMGB1 preparation used was the full-length protein purified from mammalian cells (recombinant His-tagged HMGB1 gave similar results, data not shown). The AB didomain was prepared from mammalian HMGB1 by tryptic cleavage and FPLC purification (recombinant didomain expressed in *E. coli* gave similar results). The single domains were expressed in *E. coli* (see Materials and Methods). Domain A marked as A<sub>wt</sub> contained amino-acids 1-84 and its mutant A<sub>F38A</sub> had phenylalanine at position 38 replaced with an alanine. Domain B marked as B<sub>wt</sub> contained amino-acids 85-180 and its mutants B<sub>F103A</sub> and B<sub>I122A</sub> had respectively phenylalanine at position 103 and isoleucine at position 122 replaced with an alanine, while B<sub>F+I</sub> was the double mutant. (a) Binding of full-length HMGB1. Left panel: effect of protein dilution. Labeled hcDNA concentration was 2 pM (i.e. 4 pg per sample), protein concentrations obtained by serial dilutions are indicated at the gel bottom. Right panel: competition with  $\lambda$  DNA. Labeled hcDNA concentration was 50 pM (i.e. 0.1 ng per sample); HMGB1 concentration: 150 pM in all samples except lane C; competition with  $\lambda$  DNA: 0, 10 ng, 100 ng, 1  $\mu$ g, 10  $\mu$ g, the corresponding competitor:hcDNA ratios being indicated at the gel bottom. (b) Binding of AB didomain. Left panel: effect of protein dilution. Labeled hcDNA concentration was 200 pM (i.e. 0.4 ng per sample); protein concentrations obtained by serial dilutions are indicated at the gel bottom. Right panel: competition with  $\lambda$  DNA. Labeled hcDNA concentration was 200 pM (i.e. 0.4 ng per sample); protein concentration 2 nM; competitor  $\lambda$  DNA: 0, 0.5, 5, 50, 500 ng and 5  $\mu$ g from left to right, corresponding competitor excess indicated at gel bottom. (c) Binding of domain A and mutant. Left panel: effect of protein dilution. Labeled hcDNA concentration was 50 pM (i.e. 0.1 ng per sample); protein concentrations obtained by serial dilutions as indicated at gel bottom. Right panel: competition with  $\lambda$  DNA. Labeled hcDNA concentration was 50 pM (i.e. 0.1 ng per sample); protein concentrations: A<sub>wt</sub> 2.5 nM, A<sub>F38A</sub> 10 nM; competitor  $\lambda$  DNA: 0, 10, 100, 1000 ng from left to right, corresponding competitor excess indicated at gel bottom. (d) Binding of domain B and mutants. Left panel: effect of protein dilution. Labeled hcDNA concentration was 50 pM (i.e. 0.1 ng per sample); protein concentrations obtained by serial dilutions are indicated at the gel bottom. Right panel: competition with  $\lambda$  DNA. Labeled hcDNA concentration was 50 pM (i.e. 0.1 ng per sample); protein concentration 13 nM (i.e. 3 ng per sample); competitor  $\lambda$  DNA: 0, 1, 10, 100, 1000 ng from left to right, corresponding to the indicated competitor excess. In all panels, lanes labeled C correspond to controls with no protein added.

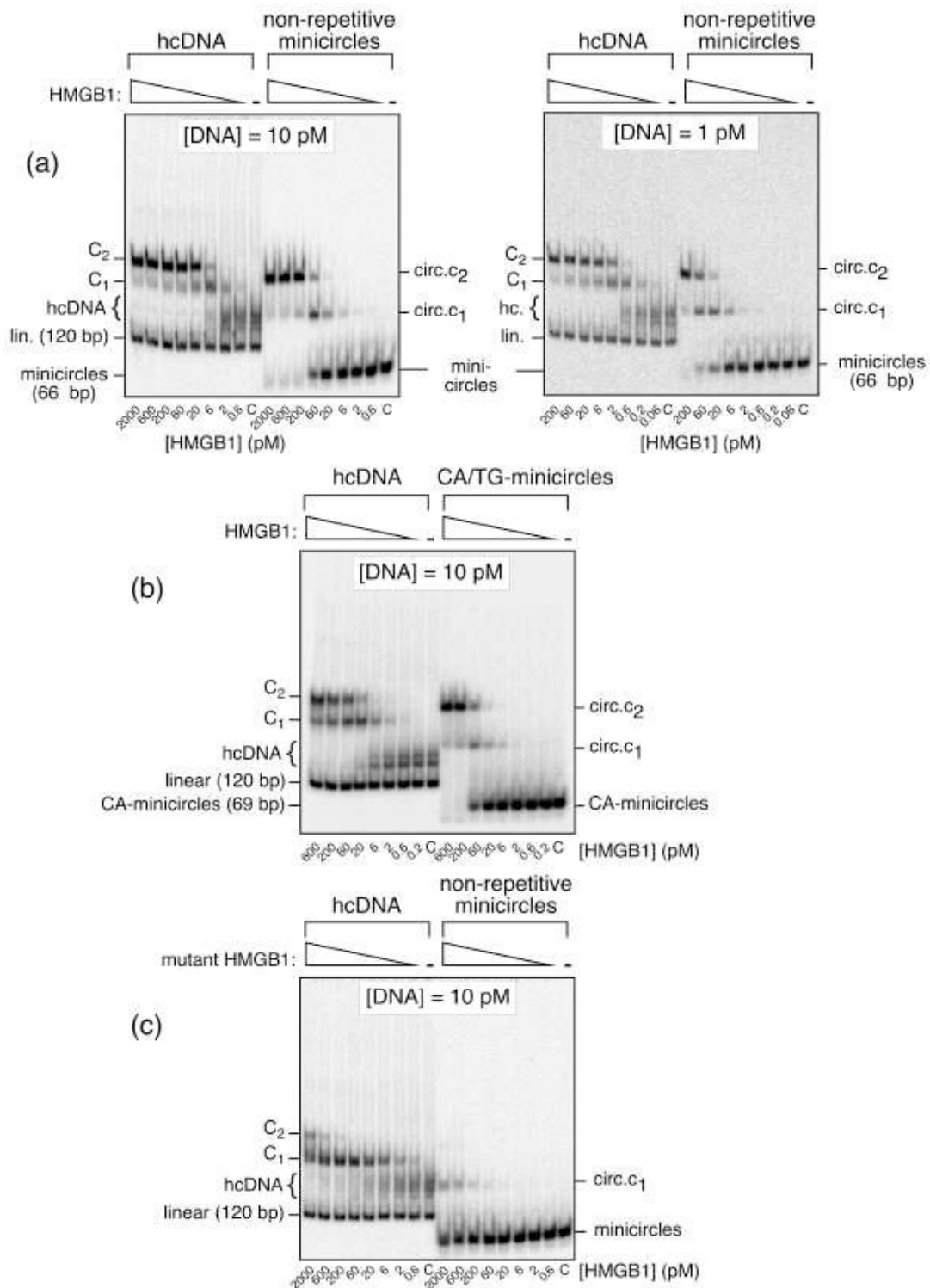


**Figure 4.** Interactions of the HMG-box domain of SRY and of protein HU with hcDNA. (a) (left panel), The HMG-box domain of human SRY (amino-acids 58-137) was expressed in *E. coli* as a (His)<sub>6</sub>-tagged protein and affinity-purified. Decreasing amounts of the protein were allowed to interact with a fixed amount of hcDNA (~50 pM), and the complexes were analyzed by electrophoresis on a polyacrylamide gel, protein concentration being indicated at the gel bottom. Lane C: control, no protein added. (a) (right panel), The interactions of SRY HMG-box with four-way junctions, hcDNA, and a linear fragment containing 4 copies of the SRY consensus recognition site were analyzed on a polyacrylamide gel. In each case, a fixed amount of labeled DNA (~50 pM) was incubated with about 3 ng (~10 nM) of protein, with 0, 50 and 500 ng of sonicated *E. coli* DNA being present as competitor (competitor excess indicated at the gel bottom). Samples labeled C are controls with no protein added. (b) (left panel), Binding of HU protein to hcDNA. Increasing amounts of HU were allowed to interact with a fixed amount of hcDNA (~50 pM) in the absence of any competitor DNA. (b) (right panel), Fixed amount of HU (1 nM) was mixed with hcDNA (~50 pM), followed by addition of increasing amounts of competitor (unlabeled) DNA. Complexes were resolved as indicated in (a).

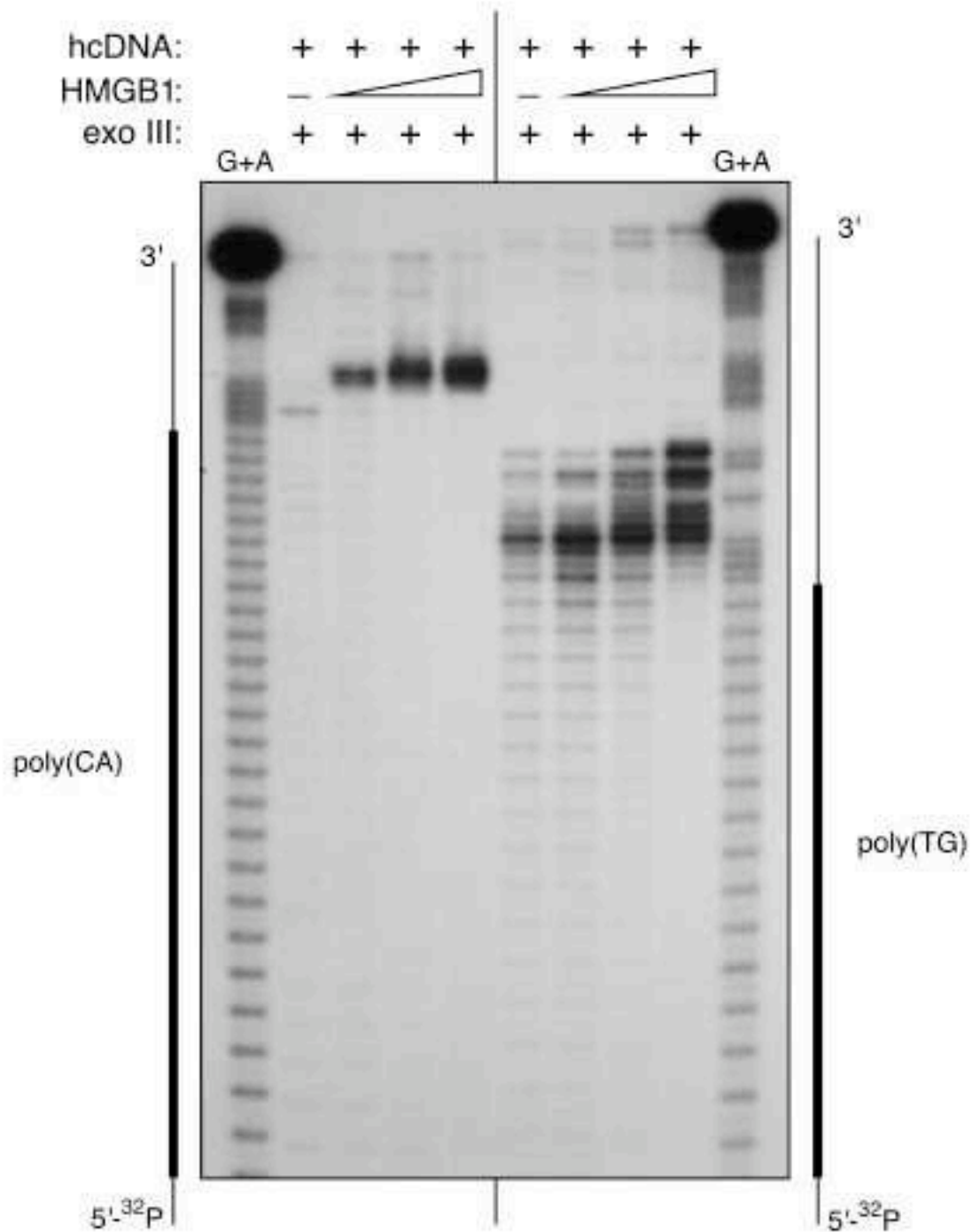


**Figure 5.** Interactions of HMGB1 with hcDNA in the presence of magnesium. Labeled hcDNA (50 pM) was incubated with decreasing amounts of protein HMGB1 in the presence of magnesium. Protein concentrations in pM are indicated at the gel bottom. After incubation, samples were analyzed on four polyacrylamide gels containing the indicated  $\text{Mg}^{++}$  concentrations. Lanes C: controls, no protein added.



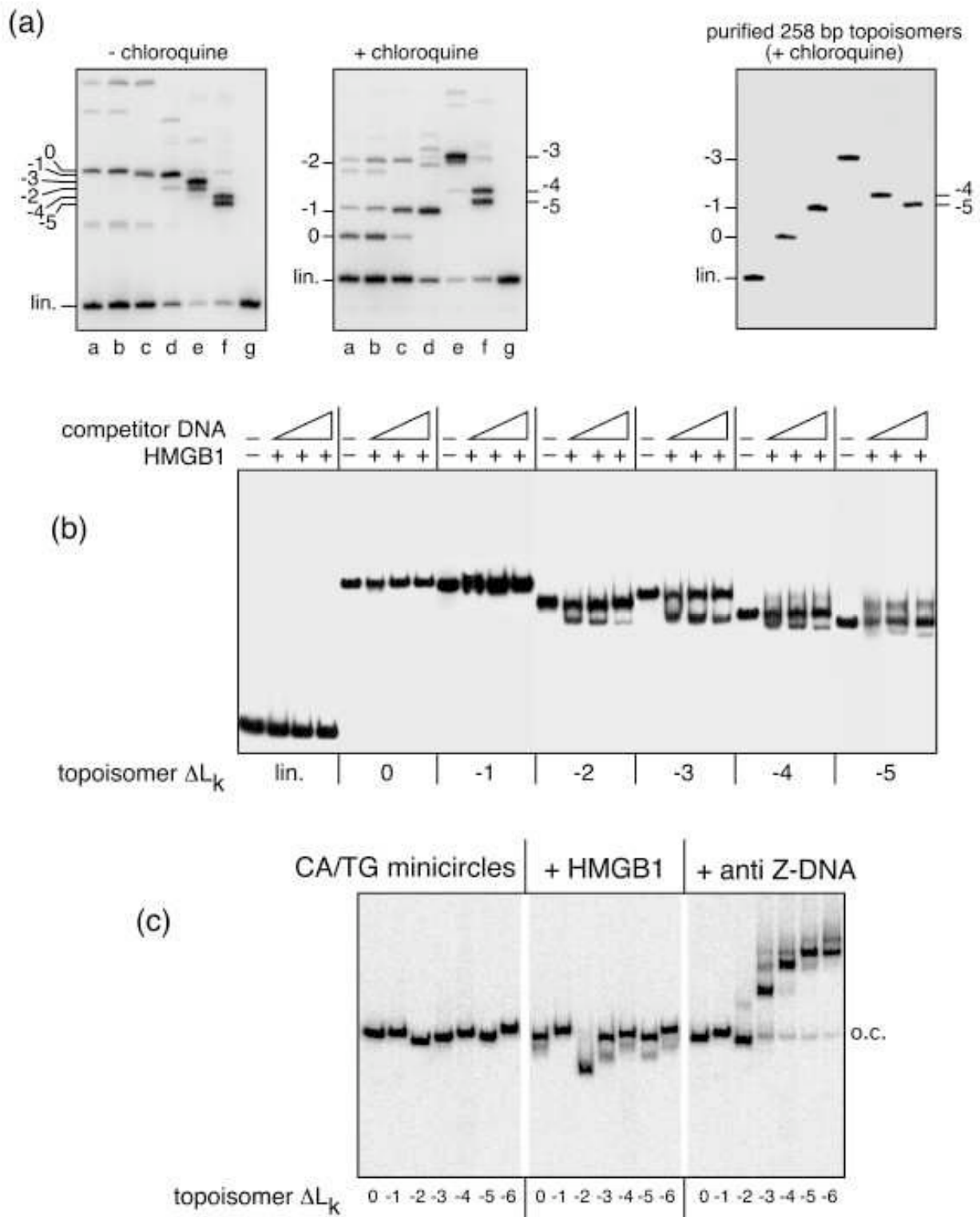


**Figure 6.** Comparison of HMGB1 and mutant HMGB1 binding to minicircles and to hcDNA. Complexes of HMGB1 or of the triple mutant protein were formed with hcDNA, 66 bp minicircles with a non-repetitive sequence, and 69 bp minicircles containing a 62 bp tract of poly(CA)·poly(TG) (see Materials and Methods). No competitor DNA was present. Protein concentrations in pM are indicated at the gel bottom. (a) Binding of HMGB1 to 66 bp non-repetitive minicircles, performed at two DNA concentrations, 10 pM and 1 pM as indicated at the top of the gels. (b) Binding of HMGB1 to 69 bp CA-minicircles. (c) Binding of mutant HMGB1 to 66 bp non-repetitive minicircles. Lanes C: controls, no protein added. Bands labelled C<sub>1</sub> and C<sub>2</sub> correspond to complexes with hcDNA, bands labelled circ.c<sub>1</sub> and circ.c<sub>2</sub> correspond to complexes with minicircles.



**Figure 7.** Exonuclease III digestion of hcDNA, and hemicatenane protection by HMGB1. hcDNA, <sup>32</sup>P-labelled at a single 5'-end (left half of the gel: *EcoRI* end; right half: *ClaI* end), was digested with exonuclease III either alone or in the presence of increasing amounts of HMGB1 (10 ng, 100 ng and 1  $\mu$ g from left to right respectively). The digestion products were analyzed by electrophoresis on a denaturing 10% polyacrylamide-urea gel and autoradiography. G+A chemical sequencing reactions were used as markers in both side lanes. The location of the poly(CA)-poly(TG) tract on the sequence is indicated on the sides of the gel. A strong pause in enzyme digestion in the presence of HMGB1 is observed at the border between the non-repetitive terminal sequences and the CA/TG repeat, immediately before the location of the hemicatenane.





**Figure 8.** Interactions of HMGB1 with individual topoisomers of a 258 bp DNA fragment. (a) Topoisomer purification. The 258 bp *Sau3AI* fragment from pBR322 was purified, 5'-end labeled, and circularized by ligation at low fragment concentration (<1 nM) in the presence of different amounts of ethidium bromide (from 0 to 2.5  $\mu\text{g}/\text{mL}$  in samples a-f, sample g containing the unligated DNA fragment) to obtain a series of circular topoisomers of the fragment. After ethidium bromide removal, the samples were analyzed by electrophoresis on 4% polyacrylamide gels, either

in the absence or in the presence of chloroquine (20  $\mu\text{g}/\text{mL}$ ). Autoradiograms of the gels are shown (a, two panels on the left), with the positions of the linear fragment (lin.) and of the six topoisomers that could be obtained, labeled 0 to -5 according to their linking number differences with the topoisomer 0 obtained in the absence of ethidium bromide, which was taken as reference. Minor bands on the gels correspond to linear oligomers and their circular forms. Individual topoisomers were cut from a preparative gel containing chloroquine and purified as described in Materials and Methods. A chloroquine-containing gel of the purified topoisomers is shown in (a), right panel (topoisomer -2 missing due to shortage of material).

(b) Interactions of HMGB1 with purified topoisomers of a labelled 258 bp fragment. The linking number differences of the topoisomers are indicated at the bottom of the gel, the topoisomer relaxed in the conditions of the experiment being taken as reference. Each sample contained  $\sim 50$  pg, i.e.  $\sim 15$  pM, of  $^{32}\text{P}$ -labeled topoisomer: For each topoisomer a series of four samples was made: one control with no protein added, and three samples containing the same amount of HeLa HMGB1 ( $\sim 600$  pg, i.e.  $\sim 1$  nM) and three different amounts of unlabeled competitor DNA (lambda phage DNA, 20 ng, 200 ng, 2  $\mu\text{g}$ , from left to right, respectively; omission of competitor DNA resulted in streaking of the samples due to low affinity of HMGB1 for all minicircles). An autoradiogram of the gel is shown.

(c) Interactions of HMGB1 with circular topoisomers containing a tract of CA/TG. Topoisomers of a  $^{32}\text{P}$ -labeled 258 bp fragment from plasmid pE10 containing a 62 bp tract of poly(CA) $\cdot$ poly(TG) were purified on a chloroquine-containing polyacrylamide gel. The topoisomers ( $\sim 50$  pg, i.e.  $\sim 15$  pM as above), in the presence of 100 ng lambda phage DNA, were incubated with HMGB1 ( $\sim 600$  pg, i.e.  $\sim 1$  nM), or with antibodies against Z-DNA, and complexes were analyzed on a polyacrylamide gel. Autoradiograms show the purified topoisomers (left), their interactions with HMGB1 (center), and the binding of anti-Z-DNA antibodies to the most supercoiled circles (right). The linking number differences of the supercoiled topoisomers relative to the relaxed circles are indicated at the bottom of the gels. o.c.: open (nicked) circles.