

## Characterization of the Stoichiometry of HMGA1/DNA Complexes

Miki Watanabe, Shuisong Ni, Amy L. Lindenberger, Junho Cho, Stuart L. Tinch and Michael A. Kennedy\*

Department of Chemistry and Biochemistry, Miami University, Oxford, OH 45056, USA

**Abstract:** High-mobility group A1 (HMGA1) non-histone chromatin architectural transcription factors regulate gene expression, embryogenesis, cell differentiation, and adaptive immune responses by binding DNA and other transcription factors. HMGA1 has also been shown to be highly over-expressed in many human cancers and is considered to be a valuable cancer biomarker. Elevated HMGA1 expression levels also make cancer cells resistant to chemotherapy. Here, HMGA1/DNA complex formation was investigated using electrophoretic mobility shift assays (EMSA). Collectively, the EMSA results indicated that full length HMGA1 mixed with DNA containing three AT-hook binding sites formed four distinct HMGA1/DNA complexes ranging in stoichiometry from 1:2 to 3:1 in HMGA1:DNA ratio. The data indicated that the distribution of complexes with different HMGA1 to DNA stoichiometries depended on the molar ratio of HMGA1 to DNA in solution, which could have significant biological implications given that HMGA1 is highly over-expressed in human cancer cells. The two naturally occurring isoforms of HMGA1, HMGA1a and HMGA1b, the latter containing an 11 amino acid deletion between the first and second AT-hooks, were observed to have slightly different DNA binding profiles. Finally, HMGA1 binding affinity to DNA was found to be influenced by the DNA A:T segment sequence context, with higher specificity be observed in HMGA1 binding to  $T_nA_n$  segments, which have two local minor groove minima on either side of the TpA step, compared to  $A_nT_n$  segments, which have a single minor groove minimum at the 3' end of the  $A_n$  run, implying AT-hook binding favors narrow minor groove structure.

**Keywords:** HMGA1, enhanceosome, AT-hook, architectural transcription factor, EMSA, DNA.

### INTRODUCTION

High-mobility group A1 (HMGA1) proteins are non-histone chromatin *architectural transcription factors* [1] found broadly among eukaryotes [2]. HMGA1 proteins, which consist of two isoforms, HMGA1a (a.k.a. HMG-I) and HMGA1b (a.k.a. HMG-Y) [3, 4], are expressed at high levels in embryonic tissues during early development [5], and at very low levels in normal adult tissues [6]. Loss of HMGA1 expression has been shown to detrimentally affect cell differentiation in embryonic stem cells [7], spermatogenesis [8], and development of type 2 diabetes and hypoglycemia in mice [9]. In chromatin function, HMGA1 proteins are thought to cause DNA destabilization associated with chromatin unfolding during DNA replication [10, 11]. Gene expression regulation is a primary normal function of HMGA1 in adults [6] with HMGA1 proteins involved in both positive and negative regulation of genes responsible for apoptosis, cell proliferation, immune response and DNA repair [12, 13]. One of the most well studied examples of HMGA1 regulation of gene expression involves the interferon- $\beta$  (IFN- $\beta$ ) gene [14, 15, 16]. IFN- $\beta$  expression is regulated by a multiple-protein/DNA complex called an *enhanceosome* [14, 17]. The IFN- $\beta$  enhanceosome, which forms in the enhancer region upstream of the gene, is

composed of multiple transcription factors including NF- $\kappa$ B, IRF, ATF2/cJun, and HMGA1. In contrast to classical transcription factors that bind specific DNA sequences, HMGA1 acts as an architectural transcription factor [1], which means that it binds a specific type of DNA structure, *i.e.* the minor groove of A:T tract DNA [18]. As architectural transcription factors, HMGA1 proteins have been suggested to play multiple roles in the formation and function of enhanceosomes including 1) induction of DNA conformational changes that increase the binding of other transcription factors to the regulatory site, 2) binding directly to other transcription factors to stabilize DNA binding, and 3) mediation of indirect interactions between proteins [12].

In addition to their role in regulating adaptive immune responses in human adults, high levels of HMGA1 expression have been reported in almost every type of human cancer [19]. High levels of HMGA1 expression have been shown to give pancreatic cancer cells resistance to chemotherapy, however, suppression of HMGA1 expression by siRNA reestablished sensitivity [20]. While the precise role that HMGA1 plays in cancer is not yet understood, HMGA1 has been suggested as a potential biomarker for tumor development and drug target for cancer therapy [21]. Several cancer therapy drugs, such as FR900482 and FL317, have been designed as competitive inhibitors of HMGA1 binding [22], however, these drugs have shown high toxicity in humans. More recently, transfection of pancreatic cancer cell lines with phosphorothioate DNA aptamers designed to

\*Address correspondence to this author at the 106 Hughes Hall, Department of Chemistry and Biochemistry, Miami University, Oxford, OH 45056, USA; Tel: 513-529-8267; Fax: 513-529-5715; E-mail: [kenned4@muohio.edu](mailto:kenned4@muohio.edu)

compete with genomic DNA for binding to HMGA1 proteins has been shown to restore sensitivity of the cancer cells to the chemotherapy agent gemcitabine [23].

Early structural studies demonstrated that HMGA1 does not adopt a conventional protein structure composed of alpha helices or beta sheets, but rather binds the minor groove of AT-rich DNA through novel DNA binding motifs referred to as "AT-hooks" [18]. AT-hooks, which are identified by the unique sequence of R-G-R-P [18, 24], adopt a crescent-shaped structure upon binding the minor groove of DNA with the arginine side chains inserting into the minor groove [18, 25]. HMGA1 is expressed as two isoforms, HMGA1a and HMGA1b, which are encoded from the same gene but HMGA1b contains a deletion of 11 amino acid residues between the first and second AT-hook as a result of the alternate splicing [3]. These isoforms have different post-translational modifications *in vivo* and are suggested to have different biological functions in tumor progression [26]. However, the difference in DNA binding activities between HMGA1a and HMGA1b has not been characterized.

Based on the present understanding of DNA binding activity of HMGA1, each AT-hook binds a minimum stretch of five or six consecutive AT base pairs. The natural substrates of HMGA1 have various lengths and sequences of AT-rich DNA [27]. In the IFN- $\beta$  enhancer region, there are four potential HMGA1 binding sites, PRDI through PRDIV [17]. Isothermal titration calorimetry experiments have shown the formation of a 1:1 complex with PRDII and also suggested HMGA1 binds to this site with two AT-hooks due to the presence of two AT-hook binding sites [28]. The promoter region of NF- $\kappa$ B has an HMGA1 binding site spanning a stretch of 19 consecutive adenines (A19) where all three AT-hooks are suggested to bind, and the positive regulatory domains I and II (PRDI and II), which each have five AT base pair segments where one AT-hook could bind [29].

The capability of HMGA1 to regulate gene expression and modulate DNA structure and stability at high concentrations could be important to unraveling its role in carcinogenesis. Since HMGA1 contains three DNA binding motifs each with different DNA affinities [30], it is plausible that several different HMGA1/DNA complexes can form with different stoichiometries when HMGA1 interacts with DNA

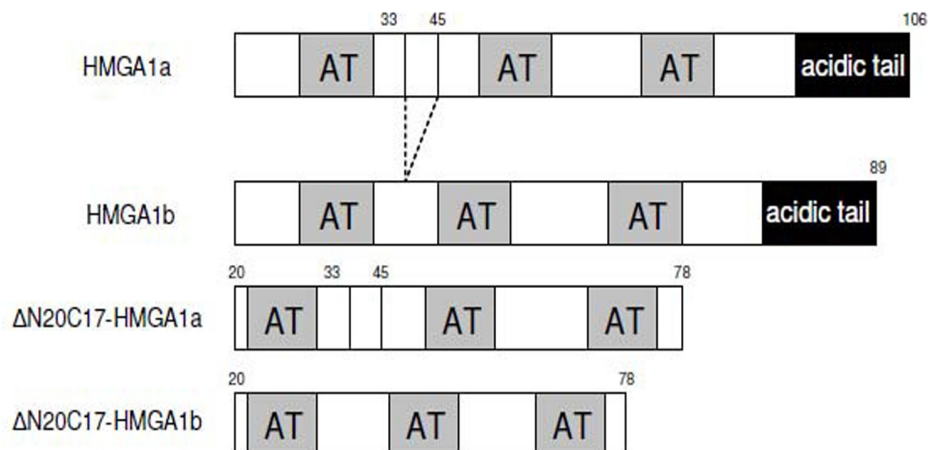
containing stretches of AT tracts that contain multiple potential AT-hook binding sites. For example, in the presence of excess protein, two or three HMGA1 proteins could potentially bind to one DNA molecule using one AT-hook from each protein. In the presence of excess DNA, HMGA1 would be capable of binding two or three DNA molecules, with each AT-hook binding a different DNA molecule.

In order to establish a better understanding of how the stoichiometry of HMGA1/DNA complexes varies with the ratio of HMGA1 to DNA in solution, we performed a series of electrophoretic mobility shift assays (EMSA) using DNA oligomers containing one, two or three AT-hook binding sites and a construct of HMGA1 that contained all three AT-hooks. The EMSA results indicated formation of four distinct complexes with stoichiometries ranging from 1:2 to 3:1 in HMGA1:DNA ratios. Our results demonstrated that the distribution of HMGA1/DNA complexes with differing stoichiometry was sensitive to the ratio of HMGA1 to DNA concentrations in solution. Furthermore, our studies revealed that the two HMGA1 isoforms, HMGA1a and HMGA1b, exhibited slightly different DNA binding behaviors, with the HMGA1b isomer exhibiting somewhat higher apparent DNA binding affinity. Finally, it was found that HMGA1 DNA binding was influenced by AT-stretch sequence context with tighter apparent binding observed in T<sub>n</sub>A<sub>n</sub> sequences, which have two local minima on either side of the TpA step, in comparison to A<sub>n</sub>T<sub>n</sub> sequences, which have only a single local minimum in minor groove width towards the 3' end of the A<sub>n</sub> run. Thus, the AT-hook binding motif appears to exhibit higher affinity and preference for narrow minor groove structures.

## MATERIALS AND METHODS

### Plasmid Construction

Deletion mutants of human HMGA1 were generated by PCR using following primers: 5'-GCCCCGCTCGAGCTASC TCCAGTTTTTGGGTCTGAAC-3' and 5'-ATCGACTCG AGCTACTCCAGTTTTTGGGTCTGCC-3'. PCR products were digested with NdeI and XhoI, purified from agarose gel by using gel extraction kit (Qiagen), and then cloned into pET-30b vector. Mutants were verified by DNA sequencing. The constructs used in the experiments are summarized in Fig. (1).



**Fig. (1).** Schematic representation of the sequences of HMGA1 isoforms and deletion mutants used in this study.

### Protein Expression and Purification

The deletion mutants,  $\Delta$ N20C17-HMGA1a and  $\Delta$ N20C17-HMGA1b, were expressed in *E. coli* BL21 (DE3). Cell culture growth was monitored by measuring the OD<sub>600</sub>. At OD<sub>600</sub> between 0.8 to 1.0, protein expression was induced by the addition of 1mM IPTG and incubated at 37°C for 4-6 hours. Deletion mutants of HMGA1a and HMGA1b were purified as described previously [31]. The proteins were further purified with Sephadex G-25 in H<sub>2</sub>O and lyophilized. The extinction coefficients of  $\Delta$ N20C17-HMGA1a and  $\Delta$ N20C17-HMGA1b at 220 nm were determined to be  $\epsilon$  220 = 63,530 M<sup>-1</sup>cm<sup>-1</sup> and  $\epsilon$  220 = 56,300 M<sup>-1</sup>cm<sup>-1</sup> by UV spectrophotometry and amino acid analysis (Purdue University). The samples were resolubilized in appropriate buffer for analysis.

### Electrophoretic Mobility Shift Assay (EMSA)

The following oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA, USA): A19 5'- CCC(A)<sub>19</sub>CC-3', A18 5'- CCC(A)<sub>18</sub>CC -3', A9T9 5'- CGAAAAAAAAAATTTTTTTTCG-3', T9A9 5'-GCTTTT TTTTAAAAAAAAAAGC-3', A15 5'- CCC (A)<sub>15</sub>CC-3', A5T5 5'-CGAAAAATTTTTTCG -3', A3T3 5'-CGAAATTT CG-3', A5 5'-GGCCCCGCGAAAAACCCGCGGG-3'. These samples were resuspended and annealed in 100mM NaCl 10mM MgCl<sub>2</sub>. Protein samples were prepared in 25 mM Tris-HCl (pH 6.5) 50 mM NaCl and the concentrations were determined based on UV 220 nm absorbance. Protein was mixed with DNA according to the ratio indicated and incubated at 4°C for 15 min prior to gel analysis. The pro-

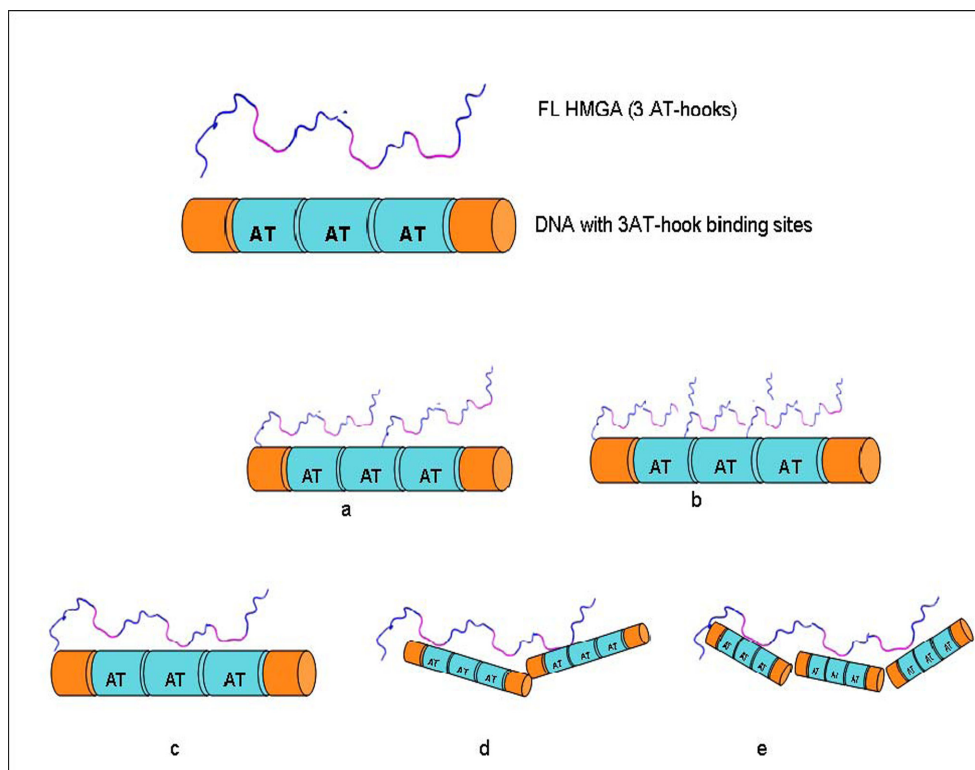
tein-DNA complexes were resolved on a 7% polyacrylamide gel and run with TAE buffer at 20 mA for 2-3 h at 4 °C. The 4-20 % precast TBE polyacrylamide gradient gel (Bio-Rad, Hercules, CA) was used for HMGA1 titration with A3T3 DNA. The gels were stained with ethidium bromide and visualized using an AlphaImager (Alpha Innotech, San Leandro, CA). The same gels were further stained with Coomassie blue and visualized using the AlphaImager.

### RESULTS AND DISCUSSION

Several HMGA1 constructs, including two different isoforms and several deletion mutants, were used in EMSA experiments (Fig. 1). Considering the nature of the HMGA1 AT-hook DNA binding motif, we anticipated that several HMGA1/DNA complexes could potentially form in the study presented here, each with a different stoichiometry, and these different possible complexes are schematically summarized in Fig. (2). Multiple approaches were used to determine the composition and stoichiometry of each band in the EMSA experiments, as described below.

#### Determination of the Stoichiometry of Binding of a Recombinant Deletion Mutant of HMGA1a to DNA Containing Various Length AT Segments

Complexes of  $\Delta$ N20C17-HMGA1a bound to DNA were studied after incubating DNA samples with increasing amounts of  $\Delta$ N20C17-HMGA1a (Fig. 3). DNA substrates contained different lengths and sequences of AT segments. The A19-DNA sequence (5'-CCCAAAAAAAAAAAAAA AAAAACC-3') is found upstream of the IL-2R $\alpha$  gene and



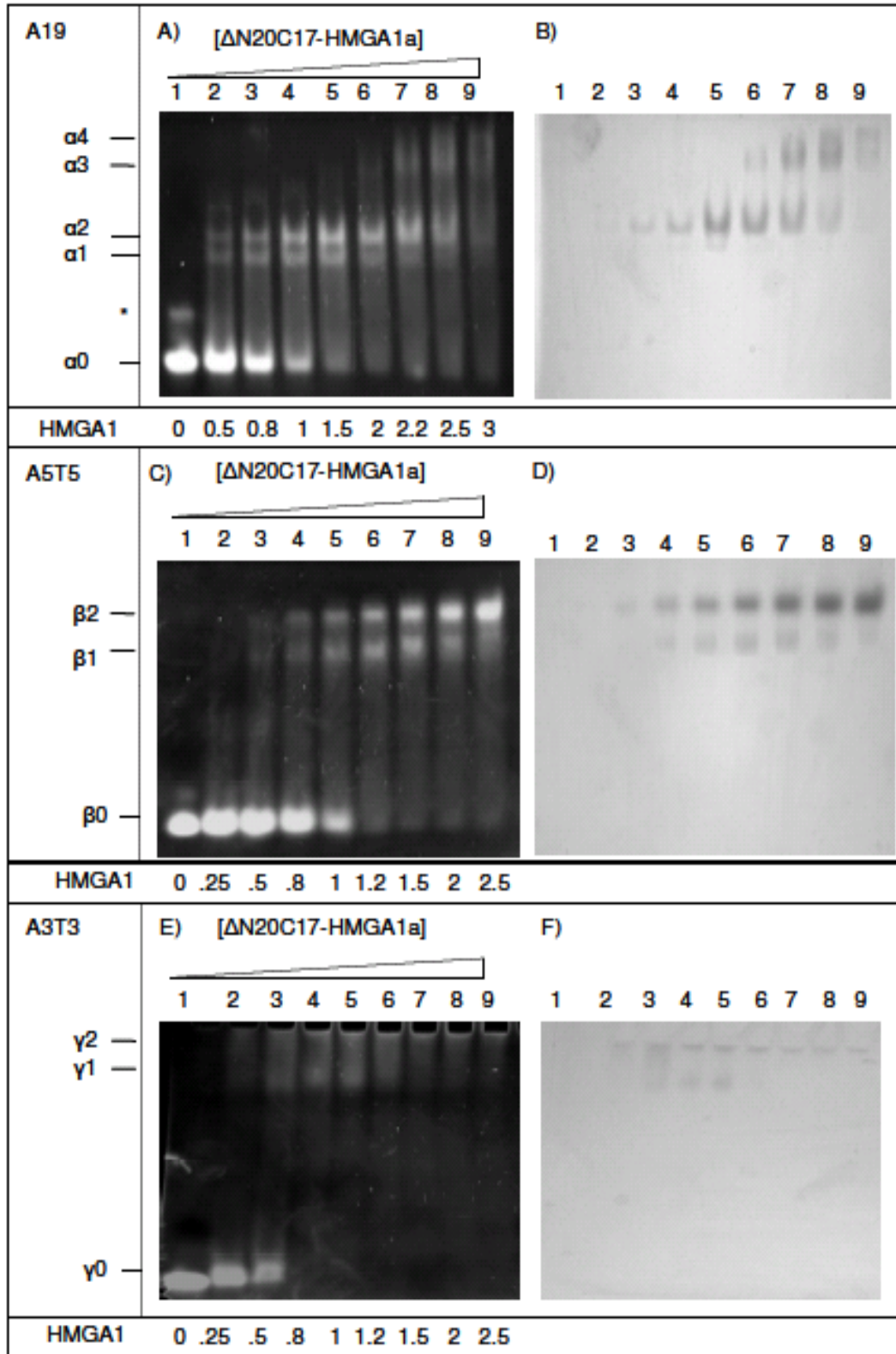
**Fig. (2).** Some potential HMGA1/DNA complexes that might form between full length HMGA1 containing three AT-hooks and DNA containing three AT-hook binding sites: (a) HMGA1:DNA (2:1), (b) HMGA1:DNA (3:1), (c) HMGA1:DNA (1:1), (d) HMGA1:DNA (1:2), (e) HMGA1:DNA (1:3).

was considered representative of a native HMGA1 substrate. In the presence of DNA containing three AT-hook binding sites, e.g. A19-DNA and A18-DNA (5'-CCCAAAAAAAAAA AAAAAAAAAACC-3'), four distinct DNA/HMGA1a complexes were detected that had different EMSA mobilities in ethidium bromide stained polyacrylamide gels (Fig. **3A** and **3B**, designated  $\alpha 1$ - $\alpha 4$ ). Two higher mobility shifted species,  $\alpha 1$  (present in Fig. **3A**, lanes 2-7) and  $\alpha 2$  (present in Fig. **3A**, lanes 2-8), were the primary species detected in titrations ranging from 1:2 to 2:1 molar ratios of  $\Delta N20C17$ -HMGA1a to A19-DNA. It also appeared that the intensity of the  $\alpha 2$  band increased with increasing  $\Delta N20C17$ -HMGA1a concentration up to about a 2:1 molar ratio of  $\Delta N20C17$ -HMGA1a to A19-DNA (Fig. **3A**, lanes 3-7) while the  $\alpha 1$  band depleted with increasing  $\Delta N20C17$ -HMGA1a concentration beyond a  $\Delta N20C17$ -HMGA1a to DNA molar ratio of greater than 1.5:1 (Fig. **3A**, lanes 5-7). In the presence of more than two-fold excess  $\Delta N20C17$ -HMGA1a, two lower mobility shifted species,  $\alpha 3$  and  $\alpha 4$ , appeared and grew stronger with increasing  $\Delta N20C17$ -HMGA1a concentration while the intensity of the  $\alpha 1$  and  $\alpha 2$  bands decreased (Fig. **3A**, lanes 6-9).

Observation of multiple bands in the EMSA following incubation of  $\Delta N20C17$ -HMGA1a with the A19-DNA sequence indicated formation of multiple complex species that differed only in the relative stoichiometry of  $\Delta N20C17$ -HMGA1a and A19-DNA. In Fig. (2), we demonstrate that, based on the number of AT-hooks present in  $\Delta N20C17$ -HMGA1a and the number of AT-hook binding sites in A19-DNA, at least five possible complexes could form. As mentioned above, we were able to detect four distinct complexes in the  $\Delta N20C17$ -HMGA1a titration with A19-DNA, indicating that we were not able to detect one of the five complexes shown in Fig. (2). Since  $\Delta N20C17$ -HMGA1a was titrated into a solution of a constant concentration of A19-DNA, the solutions in the early titration points contained an excess of A19-DNA in comparison to  $\Delta N20C17$ -HMGA1a. From the combined examination of the intensities of the bands in the ethidium bromide and Coomassie stained gels, we concluded that the 1:3  $\Delta N20C17$ -HMGA1a to A19-DNA complex was not detected and that the two higher mobility shifts,  $\alpha 1$  and  $\alpha 2$ , were the 1:2 and 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complexes, respectively. This stoichiometry assignment was based on the following observations. First of all, the  $\alpha 2$  complex was the predominant species, and presumably the most favorable complex, in solution across most of the titration. We concluded that the  $\alpha 2$  species, putatively the 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complex, was the most thermodynamically favorable complex based on both entropic considerations, *i.e.* only two molecules would have to come together to form this complex, and based on steric considerations, *i.e.* there would be the smallest steric conflicts in forming a 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complex. Secondly, the higher mobility species, *i.e.*  $\alpha 1$  in the ethidium bromide stained gel (Fig. **3A**, lane 3), would be expected to have higher overall negative charge in the complex compared to the lower mobility  $\alpha 2$  complex, and therefore the higher mobility  $\alpha 1$  complex should contain a higher DNA content compared to the lower mobility  $\alpha 2$  complex. Thirdly, inspection of lane 3 in the Coomassie stained gel (Fig. **3B**, lane 3) showed that the lower mobility  $\alpha 2$  band

complex had higher protein content (darker staining) compared to the higher mobility  $\alpha 1$  band complex. All of these observations were consistent with assignment of the  $\alpha 2$  band to a 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complex and the  $\alpha 1$  band to a 1:2  $\Delta N20C17$ -HMGA1a to A19-DNA complex. The calculated sizes of the 1:2 and 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complexes are 39.5 kDa and 23.7 kDa respectively, however, the fact that the 1:2  $\Delta N20C17$ -HMGA1a to A19-DNA complex would have a greater overall negative charge compared to the 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complex resulted in higher mobility shift in the native PAGE experiments. The mobility of the  $\Delta N20C17$ -HMGA1a to A19-DNA complex was therefore dominated by the overall charge of the complex rather than the size of the complex. Therefore, from all of these considerations, it was concluded that the  $\alpha 1$  band represented the 1:2  $\Delta N20C17$ -HMGA1a to A19-DNA complex (Fig. **2D**) and the  $\alpha 2$  band represented the 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complex (Fig. **2C**).

The lower mobility shifts in Figs. (**3A** and **3B**) ( $\alpha 3$  and  $\alpha 4$ ) were identified as 2:1 (Fig. **2A**) and 3:1 (Fig. **2B**)  $\Delta N20C17$ -HMGA1a to A19-DNA complexes, respectively. The predicted molecular weights of 2:1 and 3:1  $\Delta N20C17$ -HMGA1a to A19-DNA complexes are 31.6 kDa and 39.5 kDa, respectively. However, these complexes are predicted to have substantially less overall negative charge compared to the higher mobility 1:2 and 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complexes, which is the dominant factor responsible for their reduced EMSA mobility compared to the 1:2 and 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complexes. The lower overall net negative charge in the 2:1 and 3:1  $\Delta N20C17$ -HMGA1a to A19-DNA complexes results from multiple HMGA1a proteins, which have an isoelectric point of around pH 10, binding to a single A19-DNA molecule. Formation of multiple  $\Delta N20C17$ -HMGA1a to A19-DNA HMGA1a complexes in the presence of excess HMGA1a was verified by comparing EMSA results for  $\Delta N20C17$ -HMGA1a binding to two different DNA sequences having the same length of DNA but different numbers of AT-hook binding sites. Specifically one DNA sequence had only one AT-hook binding site, *i.e.* A5-DNA in Figs. (**3I** and **3J**), and the other sequence had three AT-hook binding sites, *i.e.* A18-DNA in Figs. (**3I** and **3J**). In the presence of A18-DNA, which is 23 base pairs long, the high mobility shifted bands,  $\alpha 1$  and  $\alpha 2$ , were observed in lanes 2-4 in Figs. (**3I** and **3J**) (as with the A19-DNA in Figs. (**3A** and **3B**) in lanes 2-8) and the lower mobility  $\alpha 3$  and  $\alpha 4$  shifted bands were also observed in lane 5 in Figs. (**3I** and **3J**) (as with the A19-DNA in Figs. (**3A** and **3B**) in lanes 7-9). However, the EMSA results for  $\Delta N20C17$ -HMGA1a to A19-DNA binding to the A5-DNA (5'-GGCCCCGCGAAAAACCCGCGGG-3'), also 23 base pairs long like the A18-DNA but with only one AT-hook binding site, showed formation of the two high mobility  $\alpha 1$  and  $\alpha 2$  bands, but no low mobility  $\alpha 3$  or  $\alpha 4$  bands were detectable in the presence of excess HMGA1a (Figs. **3I** and **3J**, lane 10). These results support the conclusion that the  $\alpha 3$  and  $\alpha 4$  bands correspond to complexes consisting of multiple  $\Delta N20C17$ -HMGA1a proteins binding to a single AT18- or AT19 DNA molecule. First of all, since the



**Fig. (3).** EMSA of  $\Delta N20C17-HMGA1a$  DNA binding activity. (A, B) EMSA of  $\Delta N20C17-HMGA1a$  with A19-DNA, (C, D) A5T5-DNA, (E, F) with A3T3-DNA, (G, H) A9T9-DNA and T9A9-DNA, (I, J) A18-DNA and A5-DNA. All gels were stained either with ethidium bromide (left) or Coomassie blue (right). The molar ratio of  $\Delta N20C17-HMGA1a$  to DNA in each lane is indicated below the gel on the left.

AT18-DNA and AT19-DNA molecules have multiple AT-hook binding sites, whereas the AT5-DNA molecule has only a single AT-hook binding site, it would be feasible for the 1:2 and 1:1  $\Delta$ N20C17-HMGA1a to AT5-DNA complexes to form but it would not be feasible to form the 2:1 and 3:1  $\Delta$ N20C17-HMGA1a to AT5-DNA complexes since the AT5-DNA only contains a single AT-hook binding site. Formation of complexes containing multiple  $\Delta$ N20C17-HMGA1a bound to a single DNA molecule with more than one AT-hook binding site in the presence of excess HMGA1a is consistent with the fact that the different AT-hooks in  $\Delta$ N20C17-HMGA1a have different DNA binding affinities. It has been shown that the second and third AT-hooks have a much higher binding affinity (nanomolar) compared to the first AT-hook (micromolar) [30]. Thus, in the presence of an excess of HMGA1, every high affinity AT-hook in  $\Delta$ N20C17-HMGA1a protein could potentially bind to its own unique DNA substrate.

### HMGA1a Binding to DNA with Two AT-Hook Binding Sites

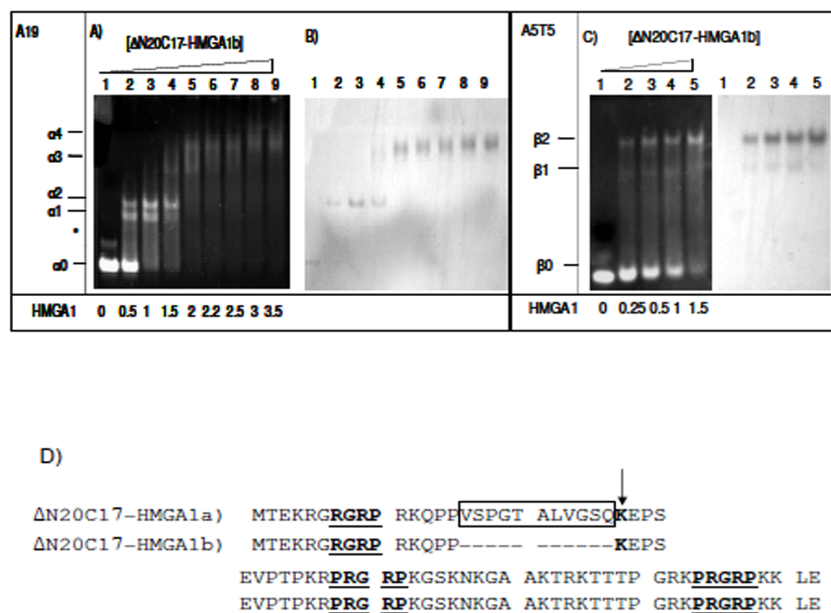
To further clarify the stoichiometries of complexes formed by  $\Delta$ N20C17-HMGA1a binding to DNA substrates, EMSA experiments were conducted using a DNA substrate containing two AT-hook binding sites, *i.e.* A5T5 (5'-CGAAAAATTTTTCG-3'). When  $\Delta$ N20C17-HMGA1a was incubated with A5T5 DNA, two  $\Delta$ N20C17-HMGA1a A5T5 species were observed corresponding to a high mobility complex ( $\beta$ 1) and a low mobility complex ( $\beta$ 2) (Figs. 3C and 3D, lanes 3-9). Both shifted bands,  $\beta$ 1 and  $\beta$ 2, were present until the concentration of HMGA1a was increased to more than two-fold excess in comparison to A5T5-DNA, after which a decrease in the intensity of the high mobility  $\beta$ 1 shifted band and an increase in the intensity of the low mobility  $\beta$ 2 shifted band was observed (Figs. 3C and 3D, lane 8-9). By a similar analysis as described above, the  $\beta$ 1 and  $\beta$ 2 shifted bands were identified as 1:2 and 1:1  $\Delta$ N20C17-HMGA1a to A5T5-DNA complexes, respectively. As above, it was expected that the higher mobility  $\beta$ 1 shifted band represented a  $\Delta$ N20C17-HMGA1a to A5T5-DNA complex that had a higher overall negative net charge in the complex compared to the lower mobility  $\beta$ 2 complex (Fig. 3C). Also, as can be seen in Fig. (3D), the lower mobility  $\beta$ 2  $\Delta$ N20C17-HMGA1a to A5T5-DNA complex had a higher relative  $\Delta$ N20C17-HMGA1a content compared to the lower mobility  $\beta$ 1 complex. These observations are both consistent with the higher mobility  $\beta$ 1 species representing a 1:2 complex  $\Delta$ N20C17-HMGA1a to A5T5-DNA complex and the lower mobility  $\beta$ 2 species representing a 1:1  $\Delta$ N20C17-HMGA1a to A5T5-DNA complex. At greater than two-fold excess of  $\Delta$ N20C17-HMGA1a to A5T5-DNA, there could be formation of complexes composed of multiple  $\Delta$ N20C17-HMGA1a proteins bound to a single A5T5-DNA molecule, however, such complexes were not detectable in these EMSA experiments, likely because they could not migrate into the gel given their increasingly positive overall charge.

### HMGA1a Binding to DNA with one AT-Hook Binding Site

To complete our analysis of  $\Delta$ N20C17-HMGA1a binding to DNA molecules containing differing numbers of AT-hook binding sites, EMSA experiments were conducted for  $\Delta$ N20C17-HMGA1a binding to a DNA substrate containing a single AT-hook binding site, namely A3T3 (5'-CGAAATTTTCG-3'). For these EMSA experiments, solutions of the complexes between  $\Delta$ N20C17-HMGA1a and A3T3-DNA had to be analyzed using a 4-20% polyacrylamide gradient gel due to the low mobility of the complexes in a 7% native gel (Fig. 3E and F). The low electrophoretic mobility of the complexes between  $\Delta$ N20C17-HMGA1a and the A3T3-DNA was consistent with diminishing overall net negative charge on the complex once  $\Delta$ N20C17-HMGA1a bound to the short ten base pair A3T3-DNA molecule. Beyond a 1:2 molar ratio of  $\Delta$ N20C17-HMGA1a to A3T3-DNA, all the A3T3-DNA was shifted primarily into a single high mobility shifted band,  $\gamma$ 1, presumably consisting of a 1:2  $\Delta$ N20C17-HMGA1a to A3T3-DNA molar ratio, plus a barely visible  $\gamma$ 2 band. Presumably, higher order  $\Delta$ N20C17-HMGA1a to A3T3-DNA complexes formed but failed to migrate into the gel, since at high  $\Delta$ N20C17-HMGA1a to A3T3-DNA molar ratios there was no detectable protein other than in the  $\gamma$ 1 shifted band in the Coomassie stained gel (Fig. 3F, lanes 6-9).

### HMGA1a Binding Activity Depends on AT Tract Sequence Context

The dependence of  $\Delta$ N20C17-HMGA1a binding on AT-DNA sequence context was also explored, and slightly different binding behavior of  $\Delta$ N20C17-HMGA1a was observed when two different DNA sequences having identical AT content but different AT sequences were used. For this experiment, the following two DNA sequences A9T9-DNA (5'-CGAAAAAAATTTTTCG-3') and T9A9-DNA (5'-GCTTTTTTTTAAAAAAAAGC-3'), were used (Fig. 3G and H). At a 1.5:1 molar ratio of  $\Delta$ N20C17-HMGA1a to T9A9-DNA, the  $\alpha$ 2 band (Fig. 3G, lane 9) was dominant, whereas with A9T9 and A18 DNA, both  $\alpha$ 1 and  $\alpha$ 2 bands were equally strong (Fig. 3G, lane 4). These results indicated that the DNA binding affinity of  $\Delta$ N20C17-HMGA1a could be influenced by AT-segment sequence context. It is well known that the DNA minor groove width is sensitive to DNA sequence [32], with self complementary sequences such as AAAAATTTTT reaching a local minimum in minor groove width at the central ApT step [33], whereas in sequences like TTTTAAAAAA, there is a local maximum in minor groove width at the TpA step, and two local minima on either side of the TpA step [32, 34, 35, 36, 37]. Poly  $A_n:T_n$  sequences, like AAAAAAAAAA:TTTTTTTTTT exhibit a narrow minor groove structure at the 3' end of the  $A_n$  stretch, however, the minor groove width does not reach the narrow magnitudes observed either in  $A_nT_n$  or  $T_nA_n$  sequences. Consequently, local variations in minor groove width appear to influence the AT-hook binding affinity and, specifically, the data suggest that the AT-hook motif has a higher affinity and specificity for narrow minor groove structure since the sequence containing two narrow minor groove



**Fig. (4).** EMSA of  $\Delta N20C17$ -HMGA1b with A19-DNA (A, B), and A5T5-DNA (C). Amino acid sequence of HMGA1a and HMGA1b deletion mutants (D). The AT-hook sequences are underlined and labeled. In  $\Delta N20C17$ -HMGA1a, the insert sequence is indicated, and in  $\Delta N20C17$ -HMGA1b, the deletion of the insertion region (boxed) results in a lysine residue (arrow) being close to first AT-hook (underlined). The molar ratio of  $\Delta N20C17$ -HMGA1b to DNA in each lane is indicated below the gel on the left.

regions, T9A9-DNA, exhibited tighter HMGA1 binding compared to a sequence with a single and less narrow minor groove region, *i.e.* A9T9-DNA.

#### Comparison of HMGA1a and HMGA1b DNA Binding Activities

EMSA experiments were repeated using the HMGA1b isoform in order to compare with the DNA binding activity of HMGA1a (Fig. 4). In the presence of A19-DNA, four different  $\Delta N20C17$ -HMGA1b/A19-DNA complexes were observed. When the molar ratio of  $\Delta N20C17$ -HMGA1b to DNA was 1.5:1 in the solution, the 1:1 complex ( $\alpha 2$ ) (Fig. 4A, lane 4 band) was dominant compared to the amount of 1:1  $\Delta N20C17$ -HMGA1a to A19-DNA complex in solution at the same protein to DNA ratio (Fig. 3A lane 5  $\alpha 2$  band). Similar results were observed for  $\Delta N20C17$ -HMGA1b binding to the A5T5 DNA (Fig. 4C). Even though two shifts were observed for  $\Delta N20C17$ -HMGA1b binding to the A5T5-DNA, as with  $\Delta N20C17$ -HMGA1a binding, in the case of the A5T5-DNA experiments the  $\beta 1$  band was very faint and the  $\beta 2$  band was clearly dominant compared to in the  $\Delta N20C17$ -HMGA1a experiment (Fig. 3C lane 5 and Fig. 4C lane 4). In both the A19-DNA and A5T5-DNA binding assays using  $\Delta N20C17$ -HMGA1b, the formation of 1:1 complexes was dominant compared to the 2:1 complex.

In  $\Delta N20C17$ -HMGA1a, the first AT-hook has been shown to have the weakest DNA binding affinity of the three AT-hooks [29, 38]. This appears to be due to the lack of positively charged amino acid residues flanking the R-G-R-P sequence. However, deletion of 11 amino acids in the

insertion region of HMGA1b places the lysine residue closer to the first R-G-R-P sequence (Fig. 4D). In addition, the lack of 11 amino acids brings two AT-hooks closer together, which may result in increased DNA binding affinity of  $\Delta N20C17$ -HMGA1b. These differences, in addition to the EMSA results, suggest that HMGA1 isoforms have subtle differences in DNA binding affinity.

#### CONCLUSIONS

Historically, EMSA studies of HMGA1 binding to DNA have been complicated to interpret because of the several possible HMGA1/DNA complex species that can form, differing only in stoichiometry, and often the precise stoichiometry of all of the bands is not specified, or even discussed. In the work presented here, not only have we demonstrated that the stoichiometry of HMGA1/DNA complexes depends sensitively on their relative molar ratios in solution, with several distinct species potentially forming both *in vitro* and *in vivo* depending on the ratio of protein to DNA, but we have also attempted to identify the precise stoichiometry of every band observed in the EMSA experiments. Furthermore, we have demonstrated, for the first time, that HMGA1 DNA binding is modulated by AT:stretch sequence context, with HMGA1 exhibiting tighter binding to TpA DNA sequences compared to  $A_n:T_n$  sequences, indicating that HMGA1 has a preference for binding to narrow minor groove structures. We have also demonstrated for the first time that the two naturally occurring isoforms of HMGA1, HMGA1a and HMGA1b, exhibit slightly different DNA binding affinities, which can be explained by considering the sequence of the amino acids flanking AT-hooks.

**ABBREVIATIONS**

EMSA = electrophoretic mobility shift assay

HMGA1 = high-mobility group A1

IFN- $\beta$  = interferon- $\beta$ 

PAGE = polyacrylamide gel electrophoresis

**FUNDING**

This work was supported by start-up funds to MAK.

**ACKNOWLEDGEMENTS**

The authors would like to acknowledge the support of Miami University and the Ohio Board of Regents for support of the Ohio Eminent Scholar Laboratory. In addition, MW thanks KNW and MW for invaluable assistance.

**CONFLICT OF INTEREST**

The authors confirm that this article content has no conflicts of interest.

**REFERENCES**

- [1] Grosschedl, R.; Giese, K.; Pagel, J. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet.*, **1994**, *10*, 94-100.
- [2] Murua Escobar, H.; Soller, J. T.; Richter, A.; Meyer, B.; Winkler, S.; Bullerdiek, J.; Nolte, I. "Best friends" sharing the HMGA1 gene: comparison of the human and canine HMGA1 to orthologous other species. *J. Hered.*, **2005**, *96*, 777-781.
- [3] Friedmann, M.; Holth, L. T.; Zoghbi, H. Y.; Reeves, R. Organization, inducible-expression and chromosome localization of the human HMG-I(Y) nonhistone protein gene. *Nucleic Acids Res.*, **1993**, *21*, 4259-4267.
- [4] Bustin, M. Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem. Sci.*, **2001**, *26*, 152-153.
- [5] Chiappetta, G.; Avantaggiato, V.; Visconti, R.; Fedele, M.; Battista, S.; Trapasso, F.; Merciai, B. M.; Fidanza, V.; Giancotti, V.; Santoro, M.; Simeone, A.; Fusco, A. High level expression of the HMGI (Y) gene during embryonic development. *Oncogene*, **1996**, *13*, 2439-2446.
- [6] Bustin, M.; Reeves, R. High-mobility group proteins: architectural components that facilitate chromatin function. *Prog. Nucleic Acid Res. Mol. Biol.*, **1996**, *54*, 35-100.
- [7] Battista, S.; Pentimalli, F.; Baldassarre, G.; Fedele, M.; Fidanza, V.; Croce, C. M.; Fusco, A. Loss of Hmgal gene function affects embryonic stem cell lympho-hematopoietic differentiation. *FASEB J.*, **2003**, *17*, 1496-1498.
- [8] Chieffi, P. HMGA1 and HMGA2 protein expression in mouse spermatogenesis. *Oncogene*, **2002**, *21*, 3644-3650.
- [9] Hock, R.; Furusawa, T.; Ueda, T.; Bustin, M. HMG chromosomal proteins in development and disease. *Trends Cell Biol.*, **2007**, *17*, 72-79.
- [10] Reeves, R.; Edberg, D. D.; Li, Y. Architectural transcription factor HMGI(Y) promotes tumor progression and mesenchymal transition of human epithelial cells. *Mol. Cell Biol.* **2001**, *21*, 575-594.
- [11] Sgarra, R.; Rustighi, A.; Tessari, M. A.; Di Bernardo, J.; Altamura, S.; Fusco, A.; Manfoletti, G.; Giancotti, V. Nuclear phospho-proteins HMGA and their relationship with chromatin structure and cancer. *FEBS Lett.*, **2004**, *574*, 1-8.
- [12] Reeves, R. Molecular biology of HMGA proteins: hubs of nuclear function. *Gene*, **2001**, *277*, 63-81.
- [13] Martinez Hoyos, J.; Fedele, M.; Battista, S.; Pentimalli, F.; Kruhoffer, M.; Arra, C.; Orntoft, T. F.; Croce, C. M.; Fusco, A. Identification of the genes up- and down-regulated by the high mobility group A1 (HMGA1) proteins: tissue specificity of the HMGA1-dependent gene regulation. *Cancer Res.*, **2004**, *64*, 5728-5735.
- [14] Thanos, D.; Maniatis, T. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell*, **1995**, *83*, 1091-1100.
- [15] Yie, J.; Liang, S.; Merika, M.; Thanos, D. Intra- and intermolecular cooperative binding of high-mobility-group protein I(Y) to the beta-interferon promoter. *Mol. Cell Biol.*, **1997**, *17*, 3649-3662.
- [16] Bonnefoy, E.; Bandu, M. T.; Doly, J. Specific binding of high-mobility-group I (HMGI) protein and histone H1 to the upstream AT-rich region of the murine beta interferon promoter: HMGI protein acts as a potential antirepressor of the promoter. *Mol. Cell Biol.*, **1999**, *19*, 2803-2816.
- [17] Thanos, D.; Maniatis, T. The high mobility group protein HMG I(Y) is required for NF-[kappa]B-dependent virus induction of the human IFN-[beta]gene. *Cell*, **1992**, *27*, 777-789.
- [18] Reeves, R.; Nissen, M.S. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J. Biol. Chem.*, **1990**, *265*, 8573-8582.
- [19] Fusco, A.; Fedele, M. Roles of HMGA proteins in cancer. *Nat. Rev. Cancer*, **2007**, *7*, 899-910.
- [20] Liao, S. S.; Whang, E. HMGA1 is a molecular determinant of chemoresistance to gemcitabine in pancreatic adenocarcinoma. *Clin. Cancer Res.*, **2008**, *14*, 1470-1477.
- [21] Liao, S. S.; Rocha, F.; Matros, E.; Redston, M.; Whang, E. High mobility group AT-hook 1 (HMGA1) is an independent prognostic factor and novel therapeutic target in pancreatic adenocarcinoma. *Cancer*, **2008**, *113*, 302-314.
- [22] Beckerbauer, L.; Tepe, J. J.; Cullison, J.; Reeves, R.; Williams, R. M.FR900482 class of anti-tumor drugs cross-links oncoprotein HMGI(Y) to DNA *in vivo*. *Chem. Biol.*, **2000**, *7*, 805-812.
- [23] Watanabe, M.; Sheriff, S.; Lewis, K. B.; Tinch, S. L.; Cho, J.; Balasubramaniam, A.; Kennedy, M. A. HMGA-targeted phosphorothioate DNA aptamers increase sensitivity to gemcitabine chemotherapy in human pancreatic cancer cell lines. *Cancer Lett.*, **2012**, *315*, 18-27.
- [24] Reeves, R.; Beckerbauer, L. HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim. Biophys. Acta*, **2001**, *1519*, 13-29.
- [25] Huth, J. R.; Bewley, C. A.; Nissen, M. S.; Evans, J. N.; Reeves, R.; Gronenborn, A. M.; Clore, G. M. The solution structure of an HMGI(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.*, **1997**, *4*, 657-665.
- [26] Edberg, D. D.; Adkins, J. N.; Springer, D. L.; Reeves, R. Dynamic and differential *in vivo* modifications of the isoform HMGA1a and HMGA1b chromatin proteins. *J. Biol. Chem.*, **2005**, *280*, 8961-8973.
- [27] Manabe, T.; Katayama, T.; Tohyama, M. HMGA1a recognition candidate DNA sequences in humans. *PLoS One*, **2009**, *4*, e8004.
- [28] Panne, D.; Maniatis, T.; Harrison, S. C. An atomic model of the interferon-beta enhanceosome. *Cell*, **2007**, *129*, 1111-1123.
- [29] John, S.; Reeves, R. B.; Lin, J.-X.; Child, R.; Leiden, J. M.; Thompson, C. B.; Leonard, W. J. Regulation of cell-type-specific interleukin-2 receptor  $\alpha$ -chain gene expression: potential role of physical interactions between Elf-1, HMGI(Y), and NF- $\kappa$ B family proteins. *Mol. Cell Biol.*, **1995**, *15*, 1786-1796.
- [30] Dragan, A. I.; Liggins, J. R.; Crane-Robinson, C.; Privalov, P. L. The energetics of specific binding of AT-hooks from HMGA1 to target DNA. *J. Mol. Biol.*, **2003**, *327*, 393-411.
- [31] Reeves, R. HMGA proteins: isolation, biochemical modifications, and nucleosome interactions. *Methods Enzymol.*, **2004**, *375*, 297-322.
- [32] Chuprina, V. P.; Lipanov, A. A.; Fedoroff, O. Y.; Kim, S. G.; Kintanar, A.; Reid, B. R. Sequence effects on local DNA topology. *Proc. Natl. Acad. Sci. U. S. A.*, **1991**, *88*, 9087-9091.
- [33] Lipanov, A. A.; Skuratovskii, I. I.; Poltev, V. I.; Golovinskaia, A. G.; Chuprina, V. P. Structure of poly(dA).poly(dT) from data of x-ray diffraction, energy calculations and nuclear magnetic resonance. *Molekuliarnaia biologiiia*, **1987**, *21*, 1645-1654.
- [34] Kennedy, M. A.; Nuutero, S. T.; Davis, J. T.; Drobny, G. P.; Reid, B. R. Mobility at the TpaA cleavage site in the T3A3-containing AhaIII and PmeI restriction sequences. *Biochemistry*, **1993**, *32*, 8022-8035.
- [35] Lingbeck, J.; Kubinec, M. G.; Miller, J.; Reid, B. R.; Drobny, G. P.; Kennedy, M. A. Effect of adenine methylation on the structure and dynamics of TpaA steps in DNA: NMR structure determination



- of [d(CGAGGTTTAAACCTCG)]<sub>2</sub> and its A9-methylated derivative at 750 MHz. *Biochemistry*, **1996**, *35*, 719-734.
- [36] McAteer, K.; Ellis, P. D.; Kennedy, M. A. The effects of sequence context on base dynamics at TpA steps in DNA studied by NMR. *Nucleic Acids Res.*, **1995**, *23*, 3962-3966.
- [37] McAteer, K.; Kennedy, M. A. NMR evidence for base dynamics at all TpA steps in DNA. *J. Biomol. Struct. Dyn.*, **2000**, *17*, 1001-1009.
- [38] Aravind, L.; Landsman, D. AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res.*, **1998**, *26*, 4413-4421.

---

Received: November 05, 2011

Revised: December 09, 2011

Accepted: December 20, 2011

© Watanabe *et al.*; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.