J. Biochem. 125, 399-405 (1999)

The Stabilized Structural Array of Two HMG1/2-Boxes Endowed by a Linker Sequence between Them Is Requisite for the Effective Binding of HMG1 with DNA

Kouhei Saito,* Takeshi Kikuchi,† Hitoshi Shirakawa,* and Michiteru Yoshida*,[1]

*Department of Biological Science and Technology and †Frontier Research Center for Computer Science, Science University of Tokyo, 2641 Yamazaki, Noda, Chiba 278-8510

Received October 19, 1998; accepted November 12, 1998

High mobility group (HMG) protein 1 contains two DNA binding motifs, called HMG1/2-boxes, linked with a linker region. The functional relationships between the two boxes and the mechanism of involvement of the linker region for effective binding of HMG1 were examined. The binding analyses of truncated HMG1 peptides with DNA indicated that the structural array of two boxes stabilizes the interaction of HMG1 with DNA. The mutation analyses of the linker region suggested that the region is equipped with tolerance for the deletion of a few amino acid residues to allow appropriate binding of the two boxes with DNA, and that the basic cluster in the linker sequence is in a position to interact with DNA. The existence of tolerance for the linker sequence was found to be conserved during the evolution of HMG1 protein homologues. A structural model for array of two boxes associating with DNA minor groove was constructed on the basis of the experimental results and energy minimization. The model proposes that the DNA binding region in HMG1 covers an 18 bp DNA region and induces its bending by about 140 degrees. The linker region may function to maintain the structural array of two HMG1/2-boxes by direct interaction with DNA.

Key words: HMG1, HMG1/2-box, model building, mutagenesis, protein-DNA interaction.

High mobility group (HMG) proteins 1 and 2 are the most prevalent nonhistone chromosomal proteins in eukaryotic organisms. Several studies have implicated HMG1 and 2 in transcription, DNA replication, and cellular differentiation. However, their functional mechanisms have not been clarified. The primary sequences of HMG1 from several species have been reported, and the protein is considered to be one of the most conserved proteins (*1-3*).

The sequence of porcine HMG1 deduced from the complementary DNA comprises 214 amino acids with a molecular weight of 24,795 Da. HMG1 consists of a simple structure containing two DNA binding motifs of about 75 amino acid residues, called HMG1/2-boxes, and an acidic C-tail containing continuous run of 30 acidic amino acids (*4*). The acidic C-tail of HMG1 is essential for the stimulation of transcription (*5*) and may participate in the destabilization of chromatin structure in transcription process (*6*), although the fine mechanism remains obscure. The HMG1/2-box has been found in RNA polymerase I transcription factor UBF (*7*), human male-determining protein SRY (*8*), lymphoid enhancer binding factor LEF-1 (*9*), mitochondrial transcription factor mtTF1 (*10*), and many nuclear and mitochondrial transcription factors (*2, 11, 12*). The binding analyses of the boxes with DNA were conducted mainly for sequence-specific HMG1/2-box proteins (*12*), because the characteristics of sequence-specificity might participate in specific gene transcription. The solution structures of isolated HMG1/2-box (*13*) and the complex with DNA (*14, 15*) have been reported for sequence-specific DNA binding boxes.

HMG1 and HMG2, sequence-nonspecific binding proteins, have a preferential affinity for unusual structured DNAs such as single-stranded DNA (*16-19*), cruciform DNA (*19-21*), B-Z junction in supercoiled DNA (*22*), and anti-cancer drug cisplatin-modified DNA (*23*). They induce unwinding (*18, 24, 25*) and bending (*26, 27*) to double-stranded DNA.

The DNA-free solution structures of both the boxes of HMG1 have been reported independently (*28-30*). However, complex structures with DNA of sequence-nonspecific boxes, including those in HMG1 as well as whole DNA binding region in HMG1, have not been clarified because of the analytical difficulties derived from its sequence-nonspecific characteristics and weak binding with DNA. Therefore, the functional significance of the presence of two boxes and the mechanism for involvement of linker region between the two boxes in association with DNA are obscure.

The present studies were conducted mainly to clarify how the linker region prescribes the structural array of the two boxes. Finally, a structural model for the complex of the DNA binding region of HMG1 with DNA is presented.

MATERIALS AND METHODS

Preparation of Deletion Peptides of HMG1 Protein—

[1] To whom correspondence should be addressed. Phone: +81-471-24-1501 (ext. 4403), Fax: +81-471-25-1841, E-mail: myoshida@rs.noda.sut.ac.jp

HMG1 peptides were overexpressed in *Escherichia coli* cells using T7 RNA polymerase system (*31*). Complementary DNA coding for peptides A (amino acid residues 1-76), Al (1-87), B (88-164), Bm (85-164), Bj (85-181), AlB (1-164), AlBj (1-181), and F (1-214) from porcine HMG1 (*4*) were amplified by polymerase chain reaction (PCR), and then inserted downstream of the T7 promoter in the expression plasmid pGEM-KI (*32*). The plasmid was transformed into *E. coli* BL21(DE3). The cultured cells of 0.3 to 0.5 absorbance at 600 nm were induced with 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG). The *E. coli* cells harvested after 6 to 10 h of induction were suspended in sonication buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 2 mM dithiothreitol, 2 mM phenylmethylsulfonyl fluoride, and 200 mM NaCl) and sonicated. The cell supernatant, obtained by centrifugation at 10,000 $\times$ *g* for 40 min, was batch-adsorbed to DEAE-cellulose. The unabsorbed protein was then fractionated with ammonium sulfate. The proteins in the supernatant of 60 or 70% saturation of ammonium sulfate were fractionated by Mono S, Mono Q, or Alkyl Superose column chromatography with a Pharmacia FPLC system. The peptides thus obtained were homogeneous on tricine-SDS-16.5% polyacrylamide gel electrophoreses followed by staining with SERVA blue G (*33*).

Electrophoretic Mobility Shift Assay—Respective peptides were analyzed for their DNA binding activity by an electrophoretic mobility shift assay using pBR322 forms I (supercoiled) and III (linearized by *Eco*RV digestion) DNA. Each peptide and 0.3 μg of DNA were incubated in 20 μl of a reaction solution (10 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM EDTA, 1 mM dithiothreitol, 10% glycerol, 20 μg bovine serum albumin, and 100 mM NaCl) at 25°C for 60 min. The complexes were applied to electrophoreses on a 1.2% agarose gel in 40 mM Tris-acetate, pH 8.0, containing 1 mM EDTA. After electrophoresis, the gel was stained with ethidium bromide.

A ^{32}P-labeled 55 bp DNA fragment with a sequence of CGG GTG AAT TCG TGT CGT GGA GCT TTG CAT ACC TTC CGA CCC AGC GAA TTC AGA G was prepared. The HMG1 deletion peptide and DNA fragment were mixed in 20 μl of the reaction solution described above, except for NaCl concentration of 50 mM, and applied to electrophoreses on an 8% polyacrylamide gel in 44.5 mM Tris-borate, pH 8.0, containing 1 mM EDTA for 3 h at 250 V.

Surface Plasmon Resonance (SPR) Measurement—A biotinylated 30 bp DNA fragment (5′-TGT ATG AAA TCT AAC AAT GCG CTC ATC GTC-3′) in TES (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 300 mM NaCl) was immobilized onto a streptavidin modified sensorchip SA (BIAcore AB) surface (*34*), resulting in the capture of approximately 1,000 RU of DNA. Each peptide at various concentrations in 10 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM EDTA, and 100 mM NaCl was injected over the immobilized DNA at a constant flow rate of 2 μl/min, and the SPR response was monitored for 600 s. The dissociation of peptide from the surface-bound complex was monitored for another 400 s at the same flow rate of the buffer. After each measurement, the DNA surface was regenerated with an injection of 8 μl of 2 M NaCl.

Data were analyzed using the BIAevalution software package, which was supplied with BIAcore. Kinetic analyses of protein-DNA interaction by the SPR measurement have been reported previously, for example, lactose operator-repressor DNA interaction (*34*), ETS1-target sequence DNA (*35*), and methionine repressor-operator complex (*36*). All data were treated according to these studies. Here, k_{ass} is the apparent association rate constant of DNA-HMG1 peptides, k_{diss} the apparent dissociation rate constant, and K_d the apparent equilibrium constant to be calculated; $K_d = k_{diss}/k_{ass}$ (*34, 36*).

Modeling the Complex of HMG1 DNA Binding Region with DNA—For modeling the structures of complexes, the PDB (Protein Data Bank of Brookhaven National Laboratory) (*37*) data for box A (1aab) (*30*) and box B (1hme) (*28*) of HMG1 and LEF-1-DNA complex (1lef) (*15*) were used. Each HMG1/2-box of HMG1 was superimposed on that of LEF-1 bound with DNA, suiting the Cα atoms of 5 amino acid residues of 9R, 10G, 15Y, 16A, 23R in 1aab and 8K, 9R, 14F, 15F, 22R in 1hme to those of 4K, 5K, 10F, 11M, 18R in 1lef which are exposed to the DNA binding surface. The box A-DNA complex was connected to the box B-DNA complex in four possible orientations. The linker region should include a part of box B because of the structural flexibility of the N-terminal part of box B. Thus the distance between the two boxes prescribed by the linker length was defined as from the C-terminal of 77Tyr residue in box A to the N-terminal of 93Ala residue in box B. The calculation was performed with InsightII/Discover3 version 97. 2 (Molecular Simulations Inc.) using amber forcefield. A 9.5 Å cut-off for non-bonded interaction and a distance-dependent dielectric function of $\varepsilon = 1 \times r$ were used. The linker peptide was initially arranged as a straight backbone placing about 10 Å apart from the box-DNA complex. The conformation of the linker peptide was calculated by 1,000 steps energy minimization and 10 ps molecular dynamics simulation at 1,000 K, 10 ps at 500 K, and 10 ps at 300 K in constraint to overlap the backbone atoms of 77Tyr and 93Ala with corresponding atoms in the two boxes. After these calculations, the linker peptide was connected with the two HMG1/2-boxes by forming peptide bonds between 76Thr and 77Tyr and between 93Ala and 94Pro, and final HMG1-DNA complex structure was energy-minimized.

RESULTS

Binding of HMG1 Peptides with DNA Observed by an Electrophoretic Mobility Shift Assay—HMG1 contains two DNA binding domains called HMG1/2-boxes of about 75 amino acid residues (box A and box B) linked by a linker (l) region of about 10 amino acid residues and an acidic C-tail of continuous 30 acidic amino acid residues connected to the box B through a joiner (j) region of about 20 amino acid residues (Fig. 1A). From the DNA binding point of view, a series of truncated products (HMG1 peptides) containing one or two HMG1/2-box(es) with or without the flanking regions was prepared. These were peptides A (1-76 amino acid residues) and B (88-164) containing a single HMG1/2-box, peptides Al (1-87), Bm (85-164), and Bj (85-181) containing all or a part of the flanking sequences, peptides AlB (1-164) and AlBj (1-181) containing two HMG1/2-boxes, and peptide F (1-214) containing the full HMG1 sequence (Fig. 1A). The DNA binding affinity of the respective peptides was primarily analyzed by an electrophoretic mobility shift assay using forms I and III DNA of

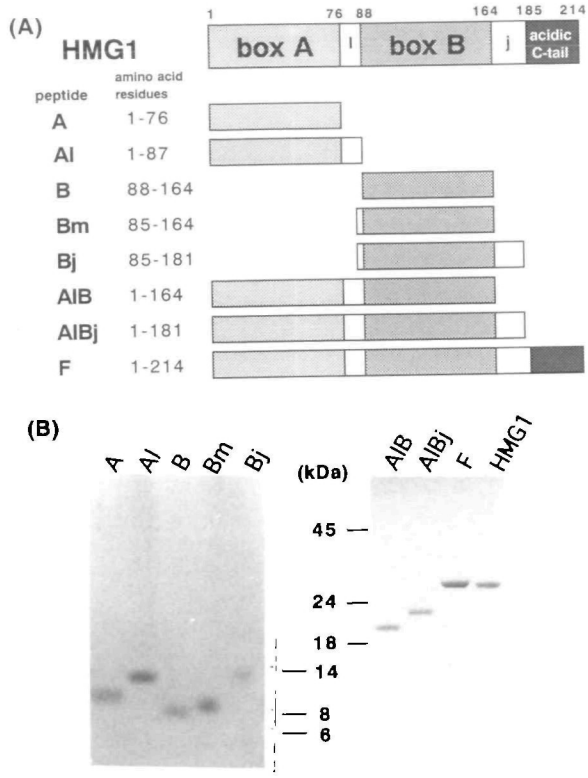

Fig. 1. **Schematic representation of HMG1 peptides.** (A) HMG1 protein consists of three domain structures; two HMG1/2-boxes of box A (1-76 amino acid residues) and box B (88-164) and an acidic C-tail (185-214). The l-region is a sequence of 77-87 residues and the j-region a sequence of 165-184 residues. (B) Electrophoretic patterns of the peptides resolved with tricine-SDS-polyacrylamide gel and stained with SERVA blue G.

plasmid pBR322 DNA. As shown in Fig. 2, all the peptides retarded the mobility of both forms I and III DNA. The peptides containing the respective flanking regions of a single HMG1/2-box (peptides Al and Bj) retarded mobility more than the corresponding HMG1/2-box alone (peptides A and B). Peptide Bm containing only 3 basic amino acid residues (KKK) at the N-terminal of peptide B showed extreme retardation compared with peptide B, indicating that the short sequence strengthens the DNA binding affinity of box B. The truncated products of acidic C-tail (peptides AlB and AlBj) showed larger band shifts than full length HMG1 (peptide F and HMG1 protein). This indicates that the acidic C-tail weakens the interaction of HMG1 with DNA. Samples containing peptides AlB and AlBj at a peptide/DNA molar ratio of 500 remained in the sample slots on the gel. Complexes at molar ratios higher than 500 migrated to the opposite electrophoretic direction from the slots in which the sample was applied (data not shown). These suggest that differences in the net charge of the peptides affect the migration of the complexes and do not simply reflect the mass of the peptides bound with DNA necessary for obtaining exact DNA binding constants.

Determination of Kinetic Constants for the Binding of HMG1 Peptides with DNA by SPR Measurements—For quantitative analysis of the DNA binding with HMG1 peptides, surface plasmon resonance (SPR) measurement using a BIAcore instrument system (BIAcore AB) was applied. Multiple sensorgrams on a linear 30 bp DNA fragment immobilized to the sensorchip surface were collected, and the DNA binding kinetics was calculated as shown in Table I. Peptides A and B gave larger dissociation constants (K_d) than other peptides. Peptide Al showed a larger association rate constant (k_{ass}) approximately 5 times that of peptide A, resulting in a decrease in K_d to one-fifth. Peptide Bm had a k_{ass} approximately 7 times larger than that of peptide B, resulting in a decrease in K_d to about one-tenth. Peptide Bj showed a lower dissociation rate constant (k_{diss}) to give a decrease in K_d to approximately one-fiftieth that of peptide B. These results suggest different roles of the l-region and the j-region with respect to the stabilization effect on DNA binding of respective HMG1/2-box. Peptides AlB and AlBj gave K_d values smaller by more than one order, together with a decrease in k_{diss} compared with peptides containing a single HMG1/2-box. Peptides AlB and AlBj gave smaller K_d values than HMG1. These results show that the structural array of two HMG1/2-boxes stabilizes the interaction with DNA, while the acidic C-tail of HMG1 weakens the binding force with DNA.

HMG1 is noticeably rich in charged amino acids; 43 lysines, 8 arginines, 19 aspartic acids, and 37 glutamic acids out of a total 214 amino acid residues (*4*). These amino acid residues have been considered to play an important role in the electrostatic interaction between HMG1 and DNA. Our examination on the interaction at various concentrations of NaCl (0-200 mM) indicated that several basic amino acid residues flanking the HMG1/2-box stabilize the DNA binding dependent on the ionic strength and that the tandem array of two HMG1/2-boxes flanked by basic amino acid residues strengthens the binding with DNA by electrostatic interaction (data not shown).

These DNA binding analyses demonstrate that a single box in HMG1 is insufficient for efficient binding with DNA even in the presence of its flanking basic region and that the sequence of peptide AlB can be regarded as the minimal DNA binding region of HMG1. Thus, the linker sequence may carry out essential roles in DNA binding.

Mutational Analysis of the Function of the l-Region in DNA Binding—To clarify the functional effects of the clustered basic amino acids in the l-region and the length of the region on the DNA binding of HMG1, several AlB mutant peptides for the l-region were constructed, as shown in Fig. 3. The K_d value for binding of each peptide with a 55 bp DNA fragment was calculated from the peptide concentration to give a half density of DNA in a lane free from peptide (Table II). Mutant 85(+0) showed a slightly decreased affinity for DNA. Deletions of 4 amino acids or fewer [mutants 85(−1) and 85(−4)] showed affinities similar to that of mutant 85(+0), whereas the affinity of 6 amino acid residues deletion mutant [mutant 85(−6)] was decreased to a level similar to that of the peptide containing a single box flanked by the l-region (peptides Al and Bm). These results suggest that the l-region is equipped with a tolerance for the deletion of a few amino acid residues to retain appropriate binding of the two boxes with DNA, and that access to DNA by either of the boxes is difficult with excessive deletion of amino acid residues. On the other hand, DNA binding was drastically reduced when amino acid residues in the basic cluster were replaced by acidic residues, while only a slight decrease was observed with small neutral amino acid residues. These

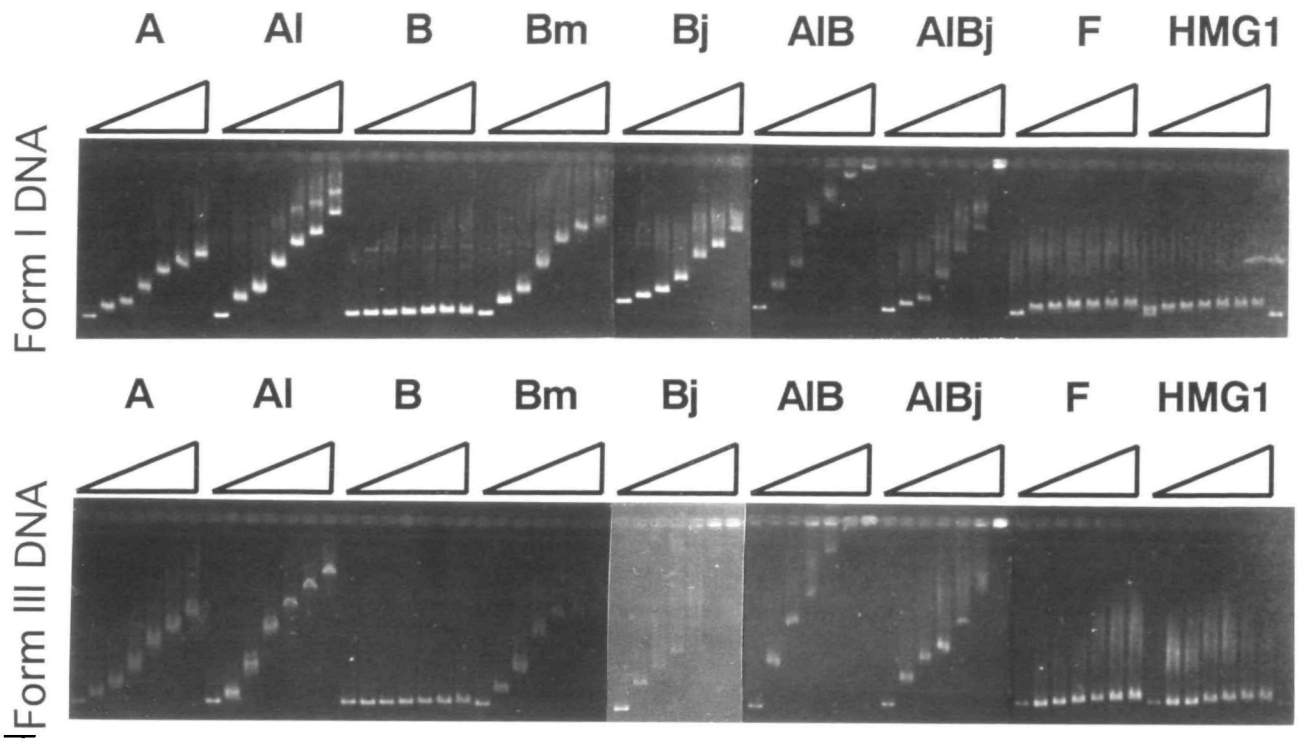

Fig. 2. **Electrophoretic mobility shift assay of HMG1 peptides.** Form I (upper panel) and form III (lower panel) pBR322 DNA were complexed with each peptide at molar ratios of 0, 50, 100, 200, 300, 400, and 500 peptides/DNA (from left to right), and then separated by 1.2% agarose gel electrophoresis.

TABLE I. **Affinity and kinetic constants for the association of the HMG1 deletion peptides with a 30 bp DNA fragment as calculated from SPR measurements.** Standard deviations are given in parentheses.

Peptide	(amino acid No.)	k_{ass} ($M^{-1}\cdot s^{-1}$)	k_{diss} (s^{-1})	K_d (M)
A	(1-76)	$4.0\,(0.3)\times10^{3}$	$2.9\,(0.3)\times10^{-1}$	$7.4\,(1.3)\times10^{-5}$
Al	(1-87)	$1.8\,(0.7)\times10^{4}$	$2.6\,(0.3)\times10^{-1}$	$1.4\,(0.1)\times10^{-5}$
B	(88-164)	$2.6\,(1.2)\times10^{3}$	$2.8\,(0.4)\times10^{-1}$	$1.3\,(0.5)\times10^{-4}$
Bm	(84-164)	$1.6\,(0.8)\times10^{4}$	$2.1\,(0.2)\times10^{-1}$	$1.4\,(0.1)\times10^{-5}$
Bj	(84-181)	$1.7\,(0.3)\times10^{4}$	$1.1\,(0.3)\times10^{-1}$	$6.7\,(1.7)\times10^{-6}$
AlB	(1-164)	$3.0\,(0.2)\times10^{4}$	$2.2\,(0.4)\times10^{-2}$	$7.5\,(1.7)\times10^{-7}$
AlBj	(1-181)	$2.1\,(0.8)\times10^{4}$	$2.0\,(0.1)\times10^{-2}$	$9.5\,(3.7)\times10^{-7}$
F	(1-214)	$1.6\,(0.2)\times10^{4}$	$6.8\,(0.7)\times10^{-3}$	$4.3\,(0.4)\times10^{-6}$
Native HMG1		$1.7\,(0.2)\times10^{4}$	$7.7\,(0.3)\times10^{-2}$	$4.6\,(0.5)\times10^{-6}$

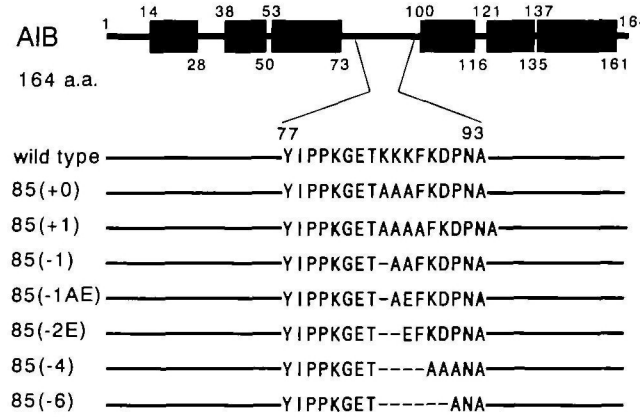

Fig. 3. **Schematic representation of the linker region mutants of HMG1 prepared for this study.** The top bar shows a simplified structure of peptide AlB illustrated by thick bars for helical structures with the numbers of their respective N- and C-terminal amino acid residues. The linker region (77-93 amino acid residues) between the two HMG1/2-boxes was substituted, deleted and substituted-deleted.

results suggest that the basic cluster lies in a position to interact with DNA.

Model Building of the Structural Array of the Two Boxes in HMG1 Bound with DNA—On the basis of the present results, an expected structural model for the array of two boxes in association with the DNA minor groove was constructed. Two assumptions were made for the model building; each sequence-nonspecific box interacts with DNA independently and in a similar fashion to sequence-specific boxes (*14, 15*). Initially, box A-DNA and box B-DNA models were constructed and connected as described in "MATERIALS AND METHODS." There were four candidates for the structural orientation of the two boxes. The linker lengths were calculated for each orientation model assuming that HMG1 covers 14 to 22 bp DNA (Fig. 4A). Models, in which less than 13 bp of DNA was covered, were excluded because of steric interference between the boxes. The linker length estimated from 77Tyr to 93Ala was 45 Å at its maximum when the backbone was completely extended. Our deletion analysis showed that the l-region has a tolerance within 4 amino acid residues to allow both boxes to interact with DNA (Table II), indicating that the linker length in a minimum of 35 Å. Although these amino acid residues do not correspond exactly to our defined borders between the l-region and the boxes, the N-terminal

TABLE II. **K_d for the association of l-region mutants of peptide AlB with 55 bp DNA as calculated from the electrophoretic mobility shift assay.**

Peptide	K_d (M)
Wild type (peptide AlB)	2.1×10^{-7}
85 (+1)	4.2×10^{-7}
85 (−1AE)	$>1.6\times10^{-6}$
85 (−2E)	$>1.6\times10^{-6}$
85 (+0)	5.1×10^{-7}
85 (−1)	5.7×10^{-7}
85 (−4)	4.8×10^{-7}
85 (−6)	8.2×10^{-7}
Peptide A	$>1.6\times10^{-6}$
Peptide Al	1.1×10^{-6}
Peptide B	$>1.6\times10^{-6}$
Peptide Bm	7.9×10^{-7}

region of box B was considered to be structurally flexible (*28*) and the flexible linker region was defined as the region between 77Tyr and 93Ala. An electrophoretic mobility shift assay of the HMG1 peptide showed that 4-5 molecules of HMG1 bind with a 74 bp DNA fragment (data not shown), indicating that one molecule of HMG1 covers a 15 to 18 bp DNA region. These limitations reduced the structural array of the two boxes to only seven possible conformations (filled circles in Fig. 4A). In addition, the results of DNA conformational change induced by the binding of HMG1 (*26*) brought further limitations. Finally, only one orientation including two structures satisfied all the limitations as discussed later. One of these structures is proposed in Fig. 4B.

DISCUSSION

The binding constants (K_d) of HMG1 with DNA obtained by SPR measurement (Table I) of 4.6×10^{-6} for native HMG1 and 4.3×10^{-6} M for recombinant HMG1 are consistent with the values obtained by the titration of tryptophan fluorescence (*38, 39*). Similar gel retardation profiles and binding constants between native and recombinant HMG1 with DNA (Fig. 2 and Table I) suggest that the native HMG1 protein is not modified and that modification, if any, does not influence the HMG1-DNA interaction.

The different K_d values for HMG1 peptides obtained by the SPR assay and the electrophoretic mobility shift assay might be resulted from differences in the principles of measurement methods, DNA fragment length and reaction buffer conditions. The electrophoretic mobility shift assay gave smaller K_d value for HMG1 than those obtained by other measurement methods (*38-40*) and by using longer DNA fragments (data not shown). The affinity of a single HMG1/2-box lacking the flanking regions was unexpectedly low (Table I). The K_d values in the presence of the flanking regions of the HMG1/2-box were decreased. The stabilization of DNA binding affinity of the HMG1/2-box in the presence of the flanking basic sequence was observed for LEF-1 (*41*). Its structural mechanism has been proposed (*15*). However, the effective binding of HMG1 required not only the presence of basic flanking region but also the structural array of two boxes. HMG1/2-box family proteins in higher eukaryotes seem to have secured stable binding ability through evolutional steps to increase the sequence-specificity of a single box and/or of the duplication of the box.

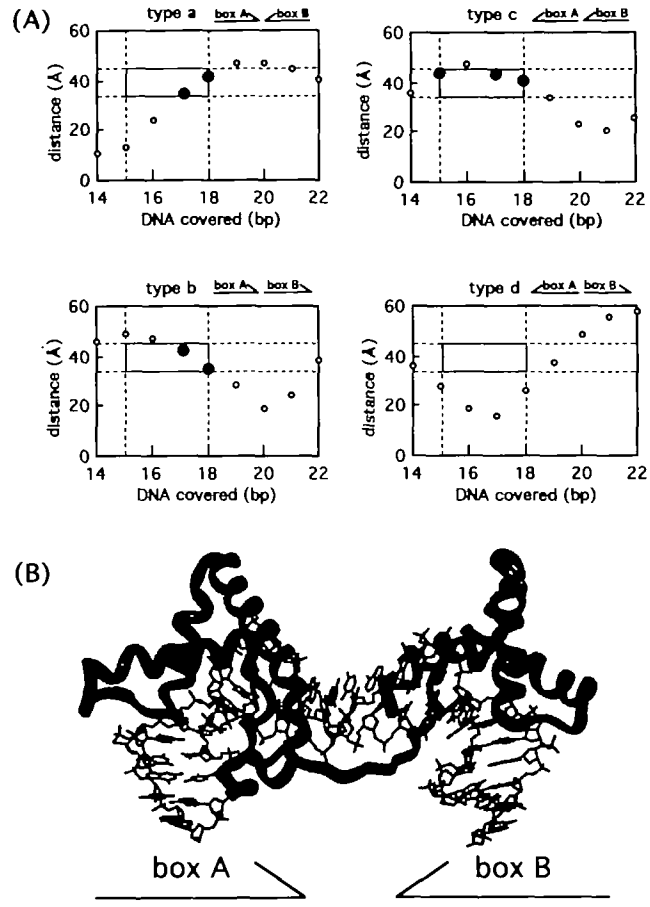

Fig. 4. **Structural array models of two HMG1/2-boxes in HMG1 on DNA.** (A) The linker distance (vertical axis) in each model was defined as the distance between the C-terminal of 77Y and the N-terminal of 93A. Four candidate arrays of the two boxes illustrated by arrows to indicate their N- to C-terminal directions were considered (types a, b, c, and d). The models, in which the DNA binding region of HMG1 covers 14-22 bp DNA region, were calculated (horizontal axis). In terms of the directions of the two boxes on DNA, only three conformations comprising seven structures were consistent with our experimental results (filled circles). (B) One of the most suitable models for the DNA binding region in HMG1 is proposed (type a model where HMG1 covers 18 bp). Two of the three conformations (types b and c) were excluded because their DNA structures were inconsistent with the results of a previous report (*26*).

The length of l-region was estimated to be between 35 and 45 Å, based on deletion experiments for the l-region (Table II). In addition, 4 or 5 molecules of peptide AlB of HMG1 bind with a 74 bp DNA fragment, showing that 15-18 bp DNA region is covered by one molecule of HMG1 protein. No structure of type d in Fig. 4A satisfied these requirements, and three orientations including seven structures remained as candidates. A 59 bp DNA fragment was ligated to generate a monomer circular in the presence of excess amount of HMG1 or 2 (*26*). The result suggested that one HMG1 molecule is able to bend DNA by more than 90 degrees. The bending angles of DNA in the models were calculated. The five structures of types b and c bent DNA by less than 60 degrees and so were excluded. Finally, one orientation model including two structures differing in the DNA spacing between the boxes by 1 bp satisfied all requirements. Figure 4B shows one of the consistent

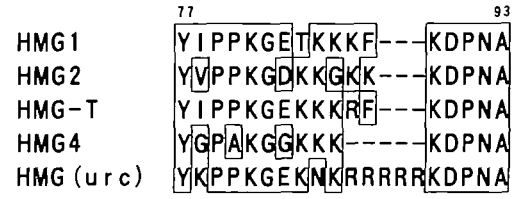

Fig. 5. **The primary sequence alignment of linker regions in HMG1 protein and its homologues.** Linker sequences of porcine HMG1 (*4*), HMG2 (*45*), trout HMG (*46*), human HMG4 (*47*), and sea urchin HMG (*48*) were aligned. Conserved amino acid residues are boxed.

models, where two boxes bind with DNA from the minor groove lying head-to-head along the strand and the l-region also goes along the minor groove. The DNA binding region in HMG1 covers 18 bp DNA region and induces bending by about 140 degrees. This model is consistent with the schematic model derived from the functional analysis of HMG2 which is highly homologues to HMG1 (*42*).

The participation in DNA binding of the linker sequence and its length between two sequence-specific DNA binding motifs of zinc fingers and POU domain was recently reported (*43, 44*). In the present study, we propose that the linker region of the sequence-nonspecific DNA binding protein HMG1 also takes part in DNA binding. HMG1 is regarded as a highly sequentially-conserved protein. The linker sequences of HMG1 and the homologues are shown in Fig. 5. It is remarkable that not only the DNA binding regions but also the linker regions between the two boxes are highly conserved. The flexibility of the linker length in HMG1 homologues is within -2 and $+3$ amino acid residues, suggesting that the enormous changes in the linker length of HMG1 homologues were excluded by natural selection due to functional insufficiency in DNA binding. The present deletion analysis strongly supported this possibility. An another interesting point is the conservation of the linker sequences among HMG1 homologues. Especially, the carboxyl-terminals of the linker sequences contain a basic cluster. The l-regions may play some roles not only in linking the two boxes but also have some unidentified functions.

The importance of length and the basic amino acid sequence in l-region between the two HMG1/2-boxes in HMG1 in effective binding with DNA was indicated in the present study. However, there remains several questions to be answered in terms of the DNA binding mechanisms, such as what is the negative function of the acidic C-tail on DNA binding and the recognition of unusual DNA structures?

REFERENCES

1. Bustin, M., Lehn, D.A., and Landsman, D. (1990) Structural features of the HMG chromosomal proteins and their genes. *Biochim. Biophys. Acta* **1049**, 231-243
2. Landsman, D. and Bustin, M. (1993) A signature for the HMG-1 box DNA binding proteins. *BioEssays* **15**, 539-546
3. Bustin, M. and Reeves, R. (1996) High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Prog. Nucleic Acid Res. Mol. Biol.* **54**, 35-100
4. Tsuda, K., Kikuchi, M., Mori, K., Waga, S., and Yoshida, M. (1988) Primary structure of non-histone protein HMG1 revealed by the nucleotide sequence. *Biochemistry* **27**, 6159-6163
5. Aizawa, S., Nishino, H., Saito, K., Kimura, K., Shirakawa, H., and Yoshida, M. (1994) Stimulation of transcription in cultured cells by high mobility group protein 1: Essential role of the acidic carboxyl-terminal region. *Biochemistry* **33**, 14690-14695
6. Ogawa, Y., Aizawa, S., Shirakawa, H., and Yoshida, M. (1995) Stimulation of transcription accompanying relaxation of chromatin structure in cells overexpressing high mobility group 1 protein. *J. Biol. Chem.* **270**, 9272-9280
7. Jantzen, H.M., Admon, A., Bell, S.P., and Tjian, R. (1990) Nuclear transcription factor hUBF contains a DNA binding motif with homology to HMG proteins. *Nature* **344**, 830-836
8. Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.M., Lovell-Badge, R., and Goodfellow, P.N. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA binding motif. *Nature* **346**, 240-244
9. Travis, A., Amsterdam, A., Belanger, C., and Grosschedl, R. (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor a-enhancer function. *Genes Dev.* **5**, 880-894
10. Parisi, M.A. and Clayton, D.A. (1991) Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science* **252**, 965-969
11. Ner, S.S. (1992) HMGs everywhere. *Curr. Biol.* **2**, 208-210
12. Grosschedl, R., Giese, K., and Pagel, J. (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet.* **10**, 94-100
13. van Houte, L.P., Chuprina, V.P., van der Wetering, M., Boelens, R., Kaptein, R., and Clevers, H. (1995) Solution structure of the sequence-specific HMG box of the lymphocyte transcriptional activator Sox-4. *J. Biol. Chem.* **270**, 30516-30524
14. Werner, M.H., Huth, J.R., Gronenborn, A.M., and Clore, G.M. (1995) Molecular basis of human 46X, Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell* **81**, 705-714
15. Love, J.J., Li, X., Case, D.A., Giese, K., Grosschedl, R., and Wright, P.E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **376**, 791-795
16. Bidney, D.L. and Reeck, G.R. (1978) Purification from cultured hepatoma cells of two nonhistone chromatin proteins with preferential affinity for single-stranded DNA. *Biochem. Biophys. Res. Commun.* **85**, 1211-1218
17. Isackson, P.J., Fishback, J.L., Bidney, D.L., and Reeck, G.R. (1979) Preferential affinity of high molecular weight high mobility group non-histone chromatin proteins for single-stranded DNA. *J. Biol. Chem.* **254**, 5569-5572
18. Yoshida, M. and Shimura, K. (1984) Unwinding of DNA by nonhistone chromosomal protein HMG(1+2) from pig thymus as determined with endonuclease. *J. Biochem.* **95**, 117-124
19. Hamada, H. and Bustin, M. (1985) Hierarchy of binding sites for chromosomal proteins HMG1 and 2 in supercoiled deoxyribonucleic acid. *Biochemistry* **24**, 1428-1433
20. Bianchi, M.E., Beltrame, M., and Paonessa, G. (1989) Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **243**, 1056-1059
21. Waga, S., Mizuno, S., and Yoshida, M. (1990) Chromosomal protein HMG1 removes the transcriptional block caused by the cruciform in supercoiled DNA. *J. Biol. Chem.* **265**, 19424-19428
22. Waga, S., Mizuno, S., and Yoshida, M. (1988) Nonhistone protein HMG1 removes the transcriptional block caused by left-handed Z-form segment in a supercoiled DNA. *Biochem. Biophys. Res. Commun.* **153**, 334-339
23. Pil, P.M. and Lippard, S.J. (1992) Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science* **256**, 234-237
24. Makiguchi, K., Chida, Y., Yoshida, M., and Shimura, K. (1984) Mg^{2+}-dependent unwinding of DNA by nonhistone chromosomal protein HMG(1+2) from pig thymus as determined by DNA melting temperature analysis. *J. Biochem.* **95**, 423-429
25. Sheflin, L.G. and Spaulding, S.W. (1989) High mobility group protein 1 preferentially conserves torsion in negatively super-

coiled DNA. *Biochemistry* **28**, 5658-5664
26. Paull, T.T., Haykinson, M.J., and Johnson, R.C. (1993) The nonspecific DNA binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structure. *Genes Dev.* **7**, 1521-1534
27. Pil, P.M., Chow, C.S., and Lippard, S.J. (1993) High-mobility-group 1 protein mediates DNA bending as determined by ring closures. *Proc. Natl. Acad. Sci. USA* **90**, 9465-9469
28. Weir, H.M., Kraulis, P.J., Hill, C.S., Raine, A.R.C., Laue, E.D., and Thomas, J.O. (1993) Structure of the HMG box motif in the B-domain of HMG1. *EMBO J.* **12**, 1311-1319
29. Read, C.M., Cary, P.D., Crane-Robinson, C., Driscoll, P.C., and Norman, D.G. (1993) Solution structure of a DNA binding domain from HMG1. *Nucleic Acids Res.* **21**, 3427-3436
30. Hardman, C.H., Broadhurst, R.W., Raine, A.R., Grasser, K.D., Thomas, J.O., and Laue, E.D. (1995) Structure of the A-domain of HMG1 and its interaction with DNA as studied by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry* **34**, 16596-16607
31. Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113-130
32. Igarashi, K. and Ishihama, A. (1991) Bipartite functional map of the *E. coli* RNA polymerase alpha subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65**, 1015-1022
33. Schagger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379
34. Bondeson, K., Frostell-Karlsson, Å., Fägerstam, L., and Magnusson, G. (1993) Lactose repressor-operator DNA interactions: Kinetic analysis by a surface plasmon resonance biosensor. *Anal. Biochem.* **214**, 245-251
35. Fisher, R.J., Fivash, M., Casas-Finet, J., Erickson, J.W., Kondoh, A., Bladen, S.V., Fisher, C., Watson, D.K., and Papas, T. (1994) Real-time DNA binding measurements of the ETS1 recombinant oncoproteins reveal significant kinetic differences between the p42 and p51 isoform. *Protein Sci.* **3**, 257-266
36. Parsons, I.D., Persson, B., Mekhalfia, A., Blackburn, G.M., and Stockley, P.G. (1995) Probing the molecular mechanism of action of co-repressor in the *E. coli* methionine repressor-operator complex using surface plasmon resonance (SPR). *Nucleic Acids Res.* **23**, 211-216
37. Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer, E.F. Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) The Protein Data Bank. A computer-based archival file for macromolecular structures. *Eur. J. Biochem.* **80**, 319-324
38. Butler, A.P., Mardian, J.K.W., and Olins, D.E. (1985) Nonhistone chromosomal protein HMG1 interaction with DNA. *J. Biol. Chem.* **260**, 10613-10620
39. Kohlstaedt, L.A. and Cole, R.D. (1994) Effect of pH on interactions between DNA and high-mobility group protein HMG1. *Biochemistry* **33**, 12702-12707
40. Carlsson, P., Waterman, M.L., and Jones, K.A. (1993) The hLEF/TCF-1a HMG protein contains a context-dependent transcriptional activation domain that induces the TCRα enhancer in T cells. *Genes Dev.* **7**, 2418-2430
41. Wagner, J.P., Quill, D.M., and Pettijohn, D.E. (1995) Increased DNA-bending activity and higher affinity DNA binding of high mobility group protein HMG-1 prepared without acids. *J. Biol. Chem.* **270**, 7394-7398
42. Shirakawa, H., Tanigawa, T., Sugiyama, S., Kobayashi, M., Terashima, T., Yoshida, K., Arai, T., and Yoshida, M. (1997) Nuclear accumulation of HMG2 protein is mediated by basic regions interspaced with a long DNA-binding sequence, and retention within the nucleus requires the acidic carboxyl terminus. *Biochemistry* **36**, 5992-5999
43. van Leeuven, H.C., Strating, M.J., Rensen, M., de Laat, W., and van del Vliet, P.C. (1997) Linker length and composition influence the flexibility of Oct-1 DNA binding. *EMBO J.* **16**, 2043-2053
44. Schwabe, J.W.R. and Rhodes, D. (1997) Linkers made to measure. *Nature Struct. Biol.* **4**, 680-683
45. Shirakawa, H., Tsuda, K., and Yoshida, M. (1990) Primary structure of non-histone chromosomal protein HMG2 revealed by the nucleotide sequence. *Biochemistry* **29**, 4419-4423
46. Pentecost, B.T., Wright, J.M., and Dixon, G.H. (1985) Isolation and sequence of cDNA clones coding for a member of the family of high mobility group proteins (HMG-T) in trout and analysis of HMG-T-mRNA's in trout tissues. *Nucleic Acids Res.* **13**, 4871-4888
47. Vaccari, T., Beltrame, M., Ferrari, S., and Bianchi, M.E. (1998) Hmg4, a new member of the Hmg1/2 gene family. *Genomics* **49**, 247-252
48. Niemeyer, C.C., Foerster-Ziober, A., and Flytzanis, C.N. (1995) Purification of a high-mobility-group 1 sea-urchin protein and cloning of cDNAs. *Gene* **164**, 211-218