Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 360 (2007) 14–19
BBRC
www.elsevier.com/locate/ybbrc

The HMGB1 acidic tail regulates HMGB1 DNA binding specificity by a unique mechanism

Qiyu Wang [a,b], Minghui Zeng [a], Wei Wang [a], Jie Tang [a,*]

[a] National Laboratory of Biomacromolecules, Center for Infection and Immunity, Institute of Biophysics, Chinese Academy of Sciences, Room 1407, 15 Datun Road, Chaoyang District, Beijing 100101, PR China
[b] School of Life Sciences, University of Science and Technology of China, Hefei 230027, PR China

Received 10 April 2007
Available online 29 May 2007

Abstract

HMGB1 is a conserved chromosomal protein composed of two DNA-binding domains and an acidic C-terminal tail. There were evidences suggesting that the C-terminal tail contributed to the DNA binding specificity of the N-terminal DNA-binding domains. However, the mechanism underlining this observation is largely unknown. Our data first confirmed the previous study with NMR that showed a direct interaction between HMGB1's C-terminal tail and its N-terminal domains. We further demonstrated that this interaction can be competed more efficiently by a DNA with four-way junction structure than by a linear double-stranded DNA. Mutations within the N-terminal region, that disrupt its binding to the C-terminal tail, abolished HMGB1's ability to distinguish the linear DNA and the four-way junction DNA. Those data suggested a unique mechanism designed by nature that utilizes a protein's negatively charged C-terminal tail to enhance its DNA-binding domain's specificity to certain structured DNAs.

Keywords: HMGB1; C-terminal tail; DNA binding specificity

High-mobility group Box 1 protein (HMGB1; also known as amphoterin or HMG1) is an abundant and highly conserved chromosomal protein. It plays a crucial role in transcriptional regulation either by remodeling chromatin and nucleosome structure or through direct interaction with transcriptional factors [1–4]. The functional importance of HMGB1 as a transcription regulator has been confirmed by phenotypes of the HMGB1 knockout mouse, which dies shortly after birth due to hypoglycemia [5]. HMGB1 also plays a role as a cytokine outside the cell, where it acts as a ligand for membrane receptor RAGE as well as Toll-like receptors 2 and 4 [6–8].

HMGB1 has two DNA-binding domains named HMG Box A, Box B (AB Boxes) and a unique carboxyl terminal domain (C-tail) that consists of a continuous array of 30 acidic amino acids [9]. HMGB1 binds without sequence specificity to the minor groove of DNAs, causing local distortions in the DNA structure [10,11]. In addition, HMGB1 associates with higher affinity to DNAs with highly bended structures, such as four-way (Holliday) junctions (4WJs) and cisplatin-modified DNAs [11]. Removal of the acidic C-tail from HMGB1 enhances its DNA-binding activity *in vitro*, but weakens its preference for bended DNAs over linear DNAs [10]. The mechanism underlining this phenotype is largely unknown. Any hypothesis explaining possible roles of the C-tail must take into consideration the very unique sequence composition of the C-tail. The 30 amino acid long C-tail domain only consists of aspartic acid and glutamic acid residues, and its sequence is highly conserved among different species. The phylogenetic conservation of this domain suggests that both the negative charges and the positions of each aspartic acid or glutamic acid residues are important for its function. There were reports indicating that this acidic tail interacts with the N-terminal part of HMGB1 protein

* Corresponding author. Fax: +86 10 6484 0932.
E-mail address: jtang@ibp.ac.cn (J. Tang).

0006-291X/$ - see front matter © 2007 Published by Elsevier Inc.
doi:10.1016/j.bbrc.2007.05.130

[12–14]. We hypothesized that the C-tail forms intra-molecular interactions with the N-terminal DNA-binding domains, and this interaction shields the DNA-binding domains from unwanted interactions with DNAs in the nucleus. This intra-molecular interaction is stronger than the binding of the AB Boxes to linear DNAs, but weaker than that to bended DNAs. In this way, the C-tail enhances HMGB1 DNA binding specificity to bended DNAs (model illustrated in Fig. 1A).

In the present study, we presented two lines of evidences to support our hypothesis. First, the C-tail bound to the AB Boxes with a K_d about 4×10^{-7} M. In a GST pull-down assay, the interaction between the C-tail and the AB Boxes could be competed more easily by a 4WJs DNA than by a linear DNA. Second, mutants in the AB Boxes disrupted its interaction with the C-tail and also abolished the preference of the full-length HMGB1 protein to a 4WJs DNA in a biotin-labeled DNA ELISA assay.

Materials and methods

Chemicals and materials. All enzymes used in the experiments were purchased from Takara (Japan) or New England BioLabs. All other reagents were purchased domestically.

Cloning, expression, and purification of proteins. The recombinant proteins corresponding to the full-length HMGB1, C-tail, AB Boxes, and their mutant versions were depicted in Fig. 1B. The cDNA of HMGB1 (encoding amino acid residues 1–215) was purchased from Invitrogen. All GST fusion constructs were cloned into the expression vector pET41a

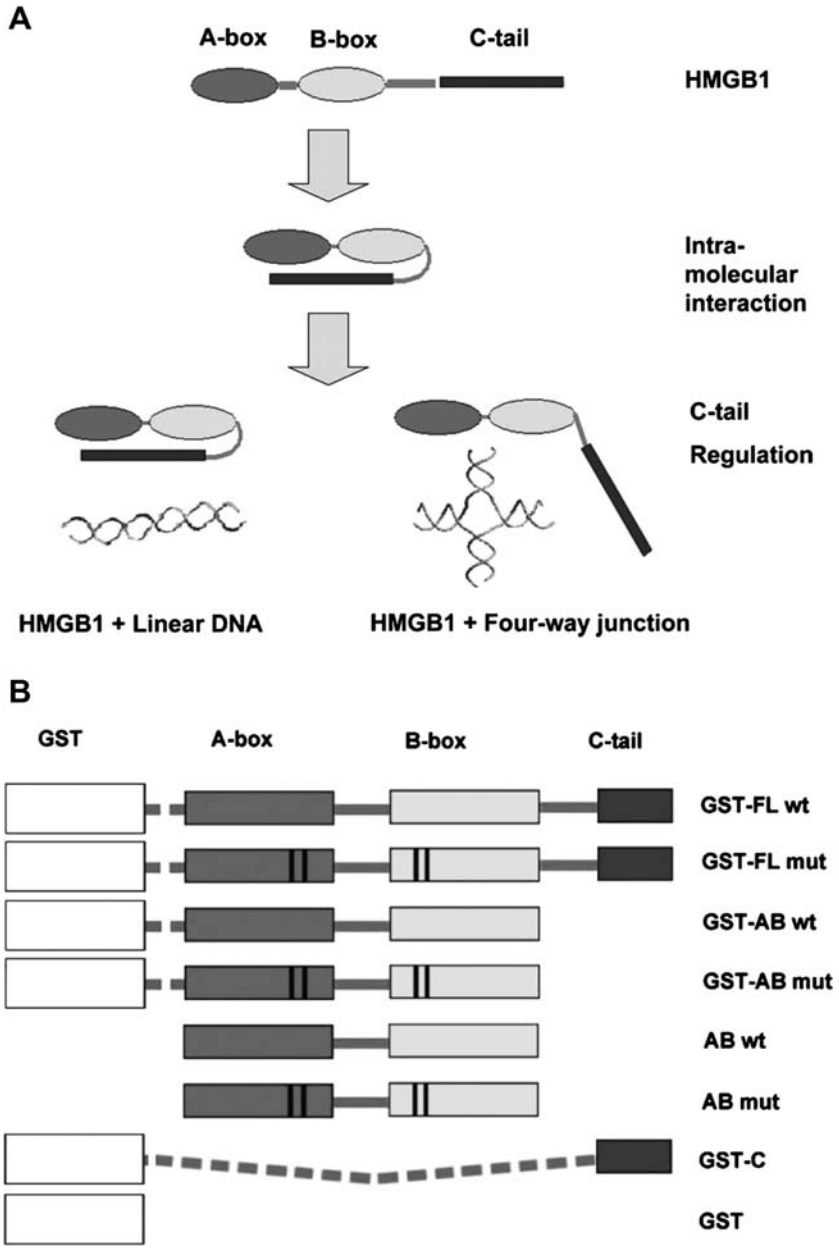

Fig. 1. (A) A proposed model in which the HMGB1 acidic tail folds back to regulate the DNA binding specificity of the N-terminal domains. (B) Schematic representation of the proteins used in this study. Black bars in Box A and B represent the mutation sites.

(Novagen) between the SpeI and XhoI sites. The wild-type and mutant AB Boxes constructs were cloned into a vector named pET41d (a pET41a variant with the GST coding sequence deleted) between the NocI and XhoI sites. ABmut or FLmut was generated by introducing four-point mutations, T76D, I78A, N92L, and A93F into ABwt or FLwt, respectively with standard PCR based mutagenesis method.

All HMGB1 proteins were expressed in *Escherichia coli* strain BL21(DE3). *E. coli* cultures were shaken at 200 rpm and 37 °C until the OD_{600} reached 0.5–0.7. Isopropyl β-D-thiogalactopyranoside (IPTG) was then added to the culture at a final concentration of 1 mM, and the protein expression was induced for an additional 4 h at 30 °C. All GST-tagged fusion proteins were purified with GST•Bind™ Purification Kits (Novagen) according to the manufacturer's protocol. All His-tagged fusion proteins were purified with His•Bind® Purification Kits (Novagen).

Surface Plasmon Resonance (SPR) analysis. The kinetics and equilibrium constant of protein–protein interactions were measured by SPR using the BIACORE3000 at 25 °C on CM5 sensor chip. Approximately 400 RU AB was captured on the flow cell. Running buffer and different concentrations GST-C (100, 200, 400, 800, and 1600 nM diluted in running buffer) were applied to the AB containing flow cell for 3 min at a flow rate of 30 μl/min. Non-specific binding to a blank flow cell was subtracted to obtain corrected sensor-grams.

Preparation of the 4WJs and linear DNA. The oligonucleotides were synthesized by Invitrogen, Beijing and purified with HPLC. The 4WJs DNA and double-stranded linear DNA were prepared by annealing equimolar amounts of the appropriate oligonucleotides in annealing buffer (10 mM Tris–HCl, pH 7.2, 50 mM NaCl, 10 mM $MgCl_2$) and slowly cooling the mixture from 94 °C to room temperature [15]. Stable formation of the 4WJs DNA was demonstrated by gel electrophoresis (data not shown). The following oligonucleotides were used to generate the 4WJs DNA: leg 1, 5′-CCC TATACCCCTGCATTGAATTCCAGTCTGATA A; leg 2, 5′-GTAGTCGTGATAGG TGCAGGGGTTATAGG; leg 3, 5′-AACAGTAGCTCTTATTCGAGCTCGCG CCCTATCACGACTA; leg 4, 5′-TTTATCAGACTGGAATTCAAGCGCGAGCTCGAATAGAG CTACTGT. A 5′-biotinylated leg 3 oligonucleotide was used to anneal with the others to generate the biotinylated 4WJs DNA.

The following oligonucleotides were used to generate the linear DNA: 25 bp-F: 5′-TTTATCAGACTGGAATTCAAGCGCG; 25 bp-R: 5′-CGC GCTTGAATTCCA GTCTGATAAA. A 5′-biotinylated 25 bp-F oligonucleotide was used to anneal with 25 bp-R to generate the biotinylated linear DNA.

GST pull-down. Fifteen micrograms of GST-C tail or GST alone was bound to glutathione beads first. Twenty micrograms of ABwt protein was incubated with the GST proteins on the beads at 4 °C for 30 min, followed by washing with PBS buffer three times. The protein complex was competed by the 4WJs or linear DNA for 30 min. Then the beads were boiled with loading buffer and proteins were separated by electrophoresis on 12% SDS–PAGE.

ELISA to detect protein and DNA interaction. Two micrograms of target protein was coated on a 96-well ELISA plate for 1 h at 37 °C. After washing with PBS buffer three times, the plate was blocked by 200 μl 1% BSA (Amresco) at 37 °C for 1 h. Then final concentration of 5, 10, 20 or 40 pmol (5 pmol/ul) biotin-labeled linear DNA was added for 1 h, with or without the non-biotinylated 4WJsDNA competition. After wash, 100 μl of 0.5 mg/L streptavidin–HRP solution (BD Biosciences Pharmingen) was added for 40 min. The *o*-phenylenediamzne solution (Amresco) was used as substrate for 3 min and the results was read at OD_{490}.

Results

The C-tail interacts with the AB Boxes

We used Surface Plasmon Resonance to directly measure the binding affinity between the purified AB Boxes (ABwt, see Fig. 1B) and the C-tail (GST-C, see Fig. 1B). The constant of dissociation between the AB Boxes and the C-tail was about 4.4×10^{-7} M (Fig. 2A). A GST pull-down assay was also designed to confirm the direct interaction between the AB Boxes and the C-tail. The GST or the GST-C was pre-bound to the glutathione beads. Purified ABwt was mixed with the GST or the GST-C on the beads and washed. The ABwt binds to the GST-C, but not the GST. The ratio between the ABwt and the GST-C is about 1:1 (Fig. 2B).

4WJs DNA competes with the C-tail in binding with the AB Boxes

When DNAs were added to the above GST pull-down assay, the 4WJs DNA could diminish the interaction

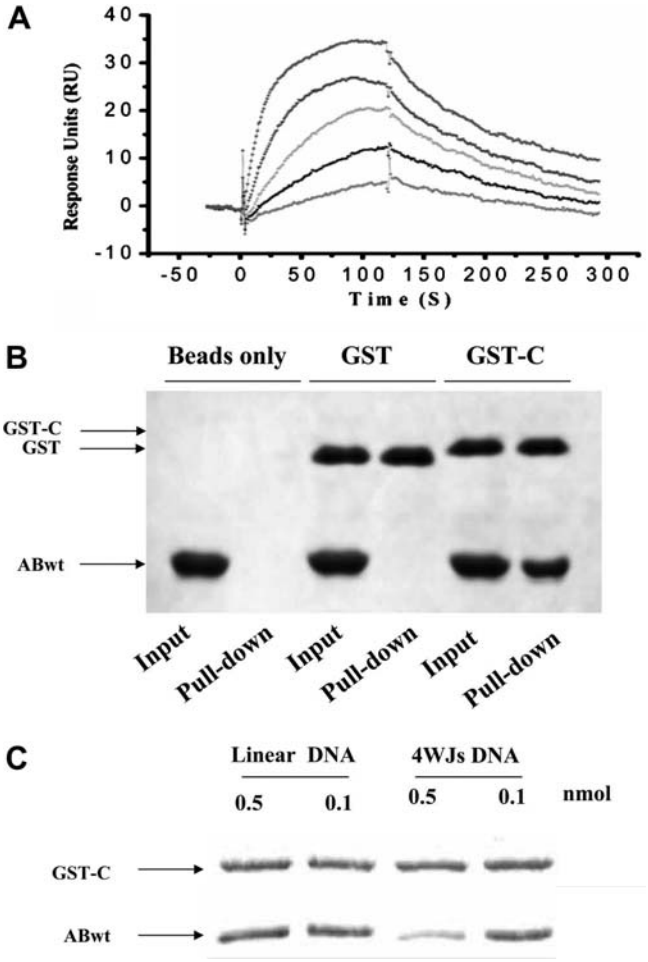

Fig. 2. The C-tail interacts with the AB boxes and this interaction can be competed by the 4WJs DNA. (A) Surface Plasmon Resonance (SPR) analysis on BIACORE3000. ABwt was coated on a CM5 sensor chip. Increasing concentration of GST-C was injected to interact with ABwt on the chip. (B) The GST control or GST-C was pre-bound to the glutathione beads. ABwt was mixed with GST or GST-C on the beads and washed three times with PBS. The protein complexes were separated by electrophoresis on 12% SDS–PAGE and visualized by Coomassie blue staining. (C) The ABwt was interacted with GST-C on the glutathione beads and washed three times with PBS. 0.1 and 0.5 nmol of the 4WJs DNA or the linear DNA was mixed with the GST-C/ABwt complex for 30 min at room temperature and washed. The protein complexes were separated by electrophoresis on 12% SDS–PAGE and visualized by Coomassie blue staining.

between the AB Boxes and the C-tail, but at the same concentration, the linear DNA could not (Fig. 2C). When the DNA concentration was further increased, both the linear DNA and the 4WJs DNA could abolish the interaction between the AB Boxes and the C-tail (data not shown). The observation suggested that there was direct competition between the C-tail and the DNAs in their binding to the AB Boxes, and the affinity of the linear DNA is lower than that of the 4WJs DNA.

The AB Boxes mutant loses its interaction with the C-tail but binds strongly with linear DNA

Previously published NMR data indicated that the C-tail may interact with T76 and I78 in Box A as well as N92, A93, I158, and R162 in Box B [12]. Four-point mutations, T76D, I78A, N92L, and A93F were introduced into one HMGB1 AB Boxes construct (ABmut, see Fig. 1B). In a GST pull-down assay, the binding of the ABmut to the GST-C is much weaker than that of the ABwt (Fig. 3A). In an ELISA assay, the GST-ABmut and GST-ABwt (see Fig. 1B) bound equally well to the biotinylated linear DNA (Fig. 3B), indicating that those mutations do not impair the AB boxes' ability to interact with DNAs.

The mutated full-length molecule loses its preference to bended DNA

The same mutations described above were introduced into a full-length HMGB1 molecule (GST-FLwt) to generate the GST-FLmut construct (see Fig. 1B). The binding of the GST-FLmut with the linear DNA is stronger than that of the GST-FLwt (Fig. 4A), but the GST-FLmut and the GST-FLwt showed a similar interaction with the 4WJs DNA (Fig. 4B). More importantly, the binding between the GST-FLwt and the linear DNA could be effectively competed by the 4WJs DNA, while the binding between the GST-FLmut and the linear DNA could not be competed by the 4WJs DNA with the same effectiveness (Fig. 4C). These data suggested that the lack of interaction between the AB boxes and the C-tail led to the loss of preference of the full-length HMGB1 to the 4WJs DNA.

Discussion

The HMGB1 contains a 30 amino acid long acidic tail that is a continuous array of aspartic acid and glutamic acid residues. This unique sequence is conserved in HMGB1 molecules of different species, ranging from

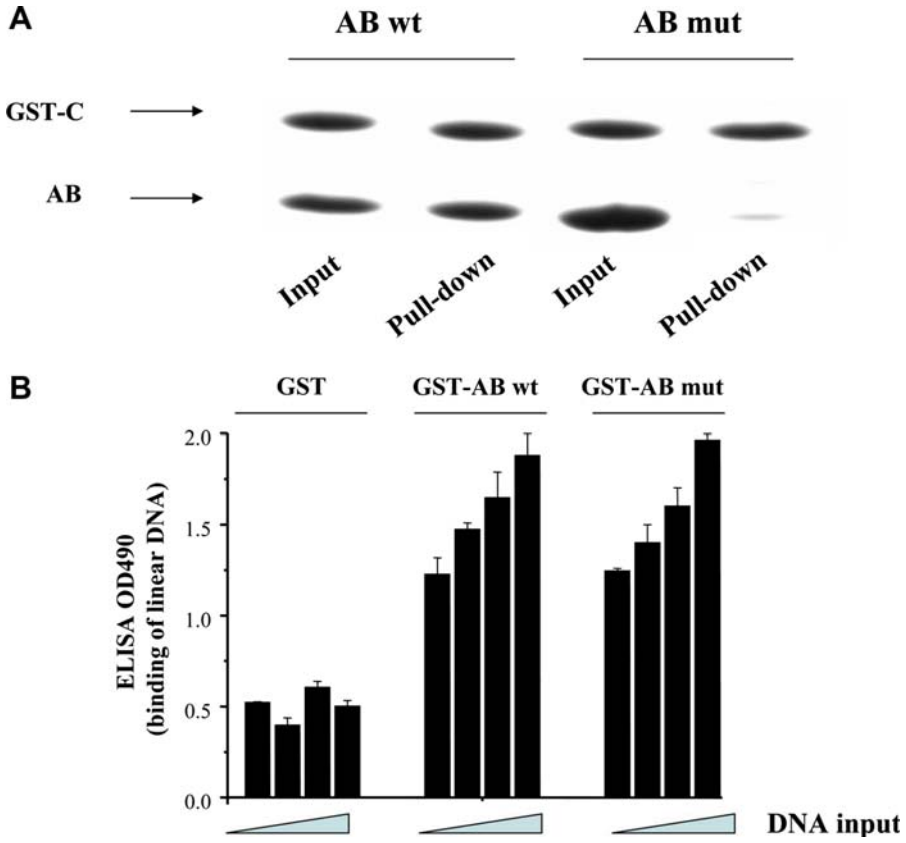

Fig. 3. The AB Box mutant no longer interacted with the C-tail but maintained its DNA binding capacity. (A) The ABwt or ABmut was incubated with pre-bound GST-C on the glutathione beads and washed. The protein complexes were separated by electrophoresis on 12% SDS–PAGE and visualized by Coomassie staining. (B) The GST, GST-ABwt or GST-ABmut was coated on a 96-well plates and incubated with the biotinylated linear DNA (5, 10, 20 or 40 pmol). The DNA–protein complex was detected by ELISA. Each error bar represents means ± SD of six measurements.

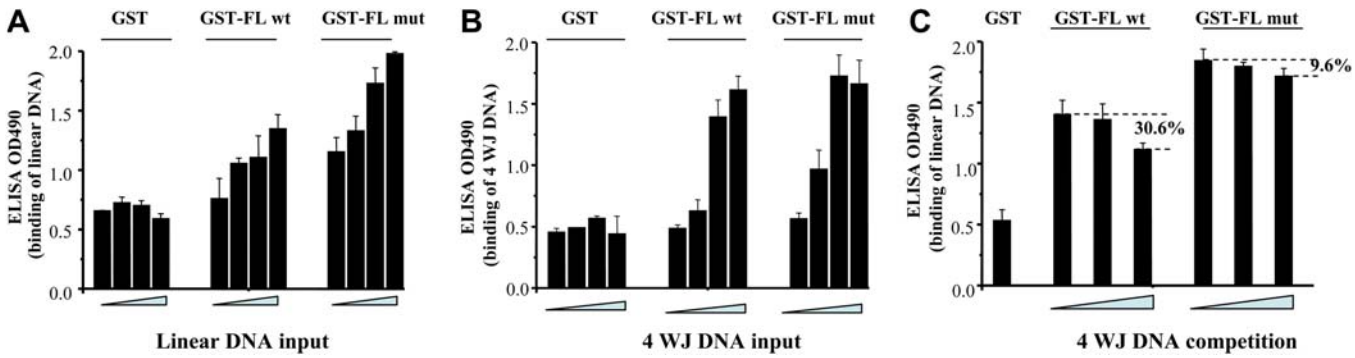

Fig. 4. The full-length mutant bound equally well to the 4WJs DNA and the linear DNA. (A) GST, GST-FLwt or GST-FLmut was coated on a 96-well plates and incubated with biotinylated linear DNA (5, 10, 20 or 40 pmol). The DNA–protein complex was detected by ELISA. (B) The GST, GST-FLwt or GST-FLmut was coated on a 96-well plates and incubated with the biotinylated 4WJs DNA (5, 10, 20 or 40 pmol). The DNA–protein complex was detected by ELISA. (C) The GST, GST-FLwt or GST-FLmut was coated on a 96-well plates and incubated with 40 pmol biotinylated linear DNA. 0, 1 or 4 times of the non-biotinylated 4WJs DNA was added as competitors. The DNA binding level was detected by ELISA. Each error bar represents means ± SD of six measurements.

human, mouse, chicken, *Xenopus* to fish. The phylogenetic conservation of the acid tail suggests a very important role it playing in HMGB1's function. This type of sequence can be found mainly in HMG family proteins. Outside of the HMG family, the similar sequence arrangement also appears in a few DNA binding proteins of certain species. However, those sequences are not conserved across different species [16–18].

We proposed here a novel model for the function of the acidic tail. In this model, the tail folds back to the N-terminal part of the protein, serving as a lid to cover the DNA binding domains. It functions as a guard only allowing the DNA with special structure to access the DNA binding domains. This feature dramatically increased the DNA binding specificity of the whole protein, so that HMGB1 will not be occupied by linear DNAs in the nucleus.

This model can explain two unique aspects of the acidic tail. First, the acidic tail binds to the same AB Boxes as the DNA do, so a continuous array of acidic residues is needed to resemble the negative charges of the DNA backbone. Second, the tail interacts with multiple residues in the AB Boxes. Since the sequence of the AB Boxes is highly conserved, the tail's sequence has to be conserved due to restrictions posted by the intra-molecular interactions. From the NMR data published in 2004 [12], we knew that residues N92 and A93 in the B Box form contacts with the C-tail. From a later published NMR structure of a complex of tandem HMG Boxes and DNA, we also knew that residues K95 and R96 are crucial for DNA binding [19]. In the B Box, the binding sites for the C-tail and the DNA are separated by only one residue. This can explain the competition observed between the C-tail and the 4WJs DNA for ABwt binding. It can also explain the fact that we can have mutations in the B box that abolished its binding to the C-tail while maintained its binding to DNAs.

Besides the intra-molecular interactions proposed here, the HMGB1 tail may also be involved in inter-molecular interactions, e.g., with the core histone H3 [20], and the TATA-binding protein (TBP) [4]. Actually, the inter- and intra-molecular interactions are not mutually exclusive. In our model, the tail forms intra-molecular interactions with the N-terminal part of the protein before the HMGB1 protein binds to DNAs, increasing the DNA binding specificity of the HMG boxes. Once the HMGB1 protein binds to its target DNA, the tail is released from the intra-molecular interactions and is free for interactions with other proteins. This type of arrangement may help to direct transcription factors to their functional sites.

Acknowledgments

This work was supported by National Natural Science Foundation of China, Grant No. 30470364. We thank Ms Yuanyuan Chen for her help on Biacore analysis.

References

[1] H. Zayed, Z. Izsvak, D. Khare, U. Heinemann, Z. Ivics, The DNA-bending protein HMGB1 is a cellular cofactor of sleeping beauty transposition, Nucleic Acids Res. 31 (2003) 2313–2322.
[2] V.S. Melvin, S.C. Roemer, M.E. Churchill, D.P. Edwards, The C-terminal extension (CTE) of the nuclear hormone receptor DNA binding domain determines interactions and functional response to the HMGB-1/-2 co-regulatory proteins, J. Biol. Chem. 277 (2002) 25115–25124.
[3] T. Imamura, H. Izumi, G. Nagatani, T. Ise, M. Nomoto, Y. Iwamoto, K. Kohno, Interaction with p53 enhances binding of cisplatin-modified DNA by high mobility group 1 protein, J. Biol. Chem. 276 (2001) 7534–7540.
[4] D. Das, W.M. Scovell, The binding interaction of HMG-1 with the TATA-binding protein/TATA complex, J. Biol. Chem. 276 (2001) 32597–32605.
[5] S. Calogero, F. Grassi, A. Aguzzi, T. Voigtlander, P. Ferrier, S. Ferrari, M.E. Bianchi, The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice, Nat. Genet. 22 (1999) 276–280.
[6] O. Hori, J. Brett, T. Slattery, R. Cao, J. Zhang, J.X. Chen, M. Nagashima, E.R. Lundh, S. Vijay, D. Nitecki, et al., The receptor for advanced glycation end products (RAGE) is a cellular binding site for

amphoterin. Mediation of neurite outgrowth and co-expression of rage and amphoterin in the developing nervous system, J. Biol. Chem. 270 (1995) 25752–25761.
[7] R. Kokkola, A. Andersson, G. Mullins, T. Ostberg, C.J. Treutiger, B. Arnold, P. Nawroth, U. Andersson, R.A. Harris, H.E. Harris, RAGE is the major receptor for the proinflammatory activity of HMGB1 in rodent macrophages, Scand. J. Immunol. 61 (2005) 1–9.
[8] M. Yu, H. Wang, A. Ding, D.T. Golenbock, E. Latz, C.J. Czura, M.J. Fenton, K.J. Tracey, H. Yang, HMGB1 signals through toll-like receptor (TLR) 4 and TLR2, Shock 26 (2006) 174–179.
[9] M.T. Lotze, K.J. Tracey, High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal, Nat. Rev. Immunol. 5 (2005) 331–342.
[10] M. Stros, J. Stokrova, J.O. Thomas, DNA looping by the HMG-box domains of HMG1 and modulation of DNA binding by the acidic C-terminal domain, Nucleic Acids Res. 22 (1994) 1044–1051.
[11] J.O. Thomas, A.A. Travers, HMG1 and 2, and related 'architectural' DNA-binding proteins, Trends Biochem. Sci. 26 (2001) 167–174.
[12] S. Knapp, S. Muller, G. Digilio, T. Bonaldi, M.E. Bianchi, G. Musco, The long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes, Biochemistry 43 (2004) 11992–11997.
[13] M.S. Thomsen, L. Franssen, D. Launholt, P. Fojan, K.D. Grasser, Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein, Biochemistry 43 (2004) 8029–8037.
[14] J. Ramstein, D. Locker, M.E. Bianchi, M. Leng, Domain-domain interactions in high mobility group 1 protein (HMG1), Eur. J. Biochem. 260 (1999) 692–700.
[15] D.A. Hill, R. Reeves, Competition between HMG-I(Y), HMG-1 and histone H1 on four-way junction DNA, Nucleic Acids Res. 25 (1997) 3523–3531.
[16] H.W. Yoon, M.C. Kim, S.Y. Lee, I. Hwang, J.D. Bahk, J.C. Hong, Y. Ishimi, M.J. Cho, Molecular cloning and functional characterization of a cDNA encoding nucleosome assembly protein 1 (NAP-1) from soybean, Mol. Gen. Genet. 249 (1995) 465–473.
[17] N. Hall, M. Karras, J.D. Raine, J.M. Carlton, T.W. Kooij, M. Berriman, L. Florens, C.S. Janssen, A. Pain, G.K. Christophides, K. James, K. Rutherford, B. Harris, D. Harris, C. Churcher, M.A. Quail, D. Ormond, J. Doggett, H.E. Trueman, J. Mendoza, S.L. Bidwell, M.A. Rajandream, D.J. Carucci, J.R. Yates 3rd, F.C. Kafatos, C.J. Janse, B. Barrell, C.M. Turner, A.P. Waters, R.E. Sinden, A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses, Science 307 (2005) 82–86.
[18] A. Kawamura, S. Koshida, H. Hijikata, A. Ohbayashi, H. Kondoh, S. Takada, Groucho-associated transcriptional repressor ripply1 is required for proper transition from the presomitic mesoderm to somites, Dev. Cell 9 (2005) 735–744.
[19] K. Stott, G.S. Tang, K.B. Lee, J.O. Thomas, Structure of a complex of tandem HMG boxes and DNA, J. Mol. Biol. 360 (2006) 90–104.
[20] T. Ueda, H. Chou, T. Kawase, H. Shirakawa, M. Yoshida, Acidic C-tail of HMGB1 is required for its target binding to nucleosome linker DNA and transcription stimulation, Biochemistry 43 (2004) 9901–9908.