

# Structural Insights into the Mechanism of Negative Regulation of Single-box High Mobility Group Proteins by the Acidic Tail Domain\*

Received for publication, June 21, 2014, and in revised form, August 21, 2014. Published, JBC Papers in Press, September 4, 2014, DOI 10.1074/jbc.M114.591115

Katherine Stott<sup>†1</sup>, Matthew Watson<sup>†1</sup>, Mark J. Bostock<sup>‡</sup>, Simon A. Mortensen<sup>§2</sup>, Andrew Travers<sup>‡</sup>, Klaus D. Grasser<sup>§</sup>, and Jean O. Thomas<sup>†3</sup>

From the <sup>†</sup>Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom and the <sup>§</sup>Department of Cell Biology and Plant Biochemistry, Biochemie-Zentrum Regensburg, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany

**Background:** HMG-box proteins bind and bend DNA, regulated by their acidic C-terminal tails.

**Results:** The acidic tails of HMG-D (*Drosophila*) and ZmHMGB1 (maize) occlude the DNA-binding regions; the strength of the interaction is phosphorylation-dependent.

**Conclusion:** Phosphorylation provides an additional mode of negative regulation.

**Significance:** Phosphorylation of the tails of insect and plant HMG-box proteins may modulate chromatin structure and accessibility.

The *Drosophila* and plant (maize) functional counterparts of the abundant vertebrate chromosomal protein HMGB1 (HMG-D and ZmHMGB1, respectively) differ from HMGB1 in having a single HMG box, as well as basic and acidic flanking regions that vary greatly in length and charge. We show that despite these variations, HMG-D and ZmHMGB1 exist in dynamic assemblies in which the basic HMG boxes and linkers associate with their intrinsically disordered, predominantly acidic, tails in a manner analogous to that observed previously for HMGB1. The DNA-binding surfaces of the boxes and linkers are occluded in “auto-inhibited” forms of the protein, which are in equilibrium with transient, more open structures that are “binding-competent.” This strongly suggests that the mechanism of auto-inhibition may be a general one. HMG-D and ZmHMGB1 differ from HMGB1 in having phosphorylation sites in their tail and linker regions. In both cases, *in vitro* phosphorylation of serine residues within the acidic tail stabilizes the assembled form, suggesting another level of regulation for interaction with DNA, chromatin, and other proteins that is not possible for the uniformly acidic (hence unphosphorylatable) tail of HMGB1.

in *Zea mays*, which bind DNA with little or no sequence specificity. The HMG box is also found in several sequence-specific transcription factors, which otherwise bear little resemblance to the non-sequence-specific HMGB proteins (1–6). The domain structure of the various abundant HMG box proteins is shown in Fig. 1 (7–10). The HMG box of *Drosophila* HMG-D is closely related to the HMGB1 B-box (40% sequence identity, 60% similarity; high structural homology), but the box has an acidic patch, and the acidic tail is markedly shorter (12 residues, including 2 serine residues, *versus* 30 consecutive acidic residues in HMGB1). The HMG box of ZmHMGB1, typical of plant chromosomal HMGB proteins, is even closer in sequence to the HMGB1 B-box than is HMG-D (48% identical, 61% similar), but also has an acidic patch (helix I), both N-terminal and C-terminal basic flanking regions, and an acidic C-terminal tail (Fig. 1), which contains several non-acidic residues (lysine, valine, asparagine, glycine, as well as serine), clustered roughly centrally in an otherwise Asp/Glu-rich region similar to the tail of HMGB1 (11).

The basic regions flanking the boxes enhance affinity for DNA (12–15), whereas the acidic tails in HMG-D (16), ZmHMGB1 (17), and HMGB1 (18) generally lower the affinity for DNA and confer selectivity for distorted DNA structures. In plant HMGB proteins, the length and positive charge of the N-terminal domain correlate with the length and negative charge of the C-terminal tail, and the two regions have been reported to interact (19).

In the structure of HMG-D bound to disulfide-cross-linked DNA (20), the basic linker appears to make contacts with the major groove via serine residues that are potentially regulated by phosphorylation; the sequence motif <sup>96</sup>KKSKK<sup>100</sup> is conserved from protozoans to human and is a substrate for protein kinase C (PKC) (21). The acidic tail contains two serine resi-

The high mobility group (HMG)<sup>4</sup> box is a DNA-binding motif that occurs in abundant chromosomal proteins such as vertebrate HMGB1 (two HMG boxes) and the functional, single-box, counterparts HMG-D in *Drosophila* and ZmHMGB1

\* This work was supported by the Biotechnology and Biological Sciences Research Council through the award of Grant BB/D002257/1 (to J. O. T.) and a grant from the Deutsche Forschungsgemeinschaft (DFG) (to K. D. G.).

⌘ Author's Choice—Final version full access.

<sup>1</sup> Both authors contributed equally to this work.

<sup>2</sup> Present address: Dept. of Biomedicine, Aarhus University, Bartholins Allé 6, 8000 Aarhus C, Denmark.

<sup>3</sup> To whom correspondence should be addressed. Tel.: 44-1223-333670; E-mail: jot1@cam.ac.uk.

<sup>4</sup> The abbreviations used are: HMG, high mobility group; HMGB, high mobility group box; HSQC, heteronuclear single quantum coherence; TOCSY, total correlation spectroscopy; PRE, paramagnetic relaxation enhancement;

MTSL, [1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidin-3-yl] methyl methanethiosulfonate; ATSL, [1-acetoxy-2,2,5,5-tetramethyl-3-pyrrolidin-3-yl] methyl methanethiosulfonate; CK2, casein kinase 2; P, phosphorylated.

# Role of the Acidic Tail of Single HMG-box Proteins

## Eukaryotic HMGB proteins

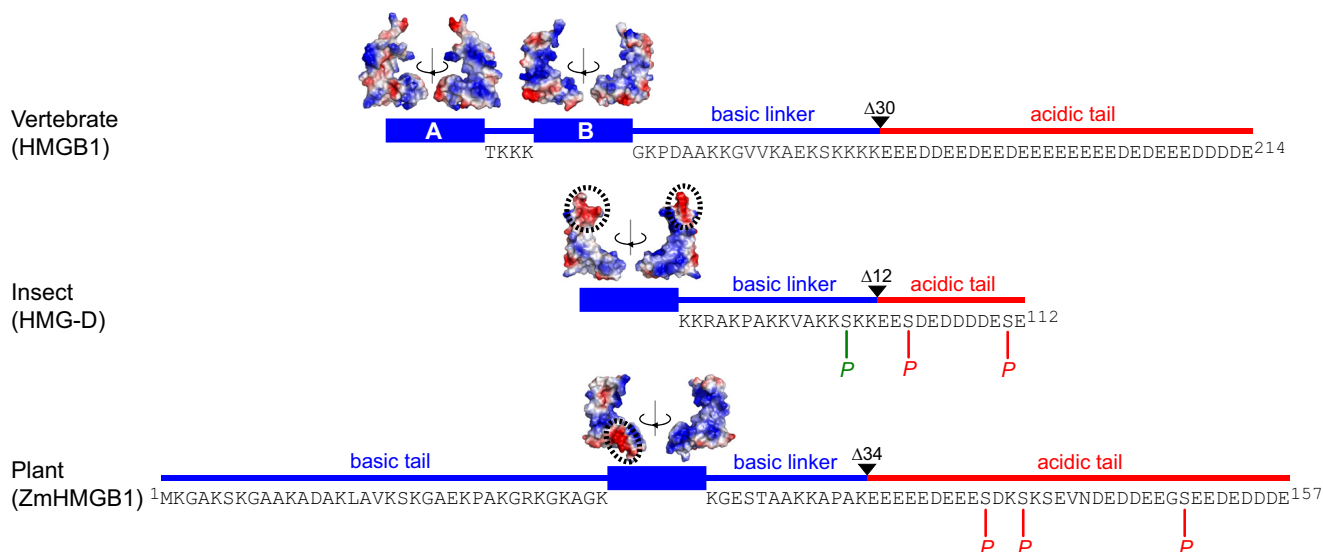


FIGURE 1. **Domain organization of eukaryotic HMGB proteins.** Sequences of linker and tail regions are shown. Truncations are indicated by  $\Delta$ , and the CK2 (red) and PKC (green) phosphorylation sites are indicated by P. Surface electrostatics for the HMG boxes (7–9) were generated in PyMOL (10); acidic patches (see the Introduction) are encircled by dotted lines. The structure of the ZmHMGB1 box (residues 35–112) was derived by homology modeling (see “Experimental Procedures”).

dues, Ser<sup>103</sup> and Ser<sup>111</sup>, that are constitutively phosphorylated by casein kinase 2 (CK2) (22); *in vitro* dephosphorylation of these residues decreases the stability of native HMG-D and reduces the selectivity of the protein for distorted DNA substrates. The interaction between the basic and acidic tail regions of ZmHMGB1 is increased upon phosphorylation (by CK2) of serine residues in the acidic tail (23–25), and the affinity of the protein for linear DNA is decreased. In addition to the effects on DNA binding and stability, phosphorylation has also been shown to affect the distribution of HMGB proteins within the nucleus in *Arabidopsis* (26) and between the nucleus and cytoplasm in the *Chironomus* HMG box proteins (closely related to HMG-D) (27).

Given the differences between HMG box proteins that exist outside the box itself (Fig. 1), it is unclear whether, like vertebrate HMGB1 (28, 29), the single-box proteins exist in a dynamic assembly in which the HMG box associates intramolecularly with the acidic tail. In HMGB1, the DNA-binding surfaces of the two boxes and linkers are occluded in an “auto-inhibited” form of the protein, which is in equilibrium with transient, more open structures that are “binding-competent.” Although the mechanism is potentially a general one for regulation of DNA binding by acidic tails, this has not been demonstrated, and the differences illustrated in Fig. 1 may play a role. For example, the acidic patch on the HMG-D box could result in a lower affinity for its short acidic tail, and phosphorylation of the tail in HMG-D and ZmHMGB1 may play a part in modulating its intramolecular interactions.

Here we show that in both HMG-D and ZmHMGB1, the acidic tail is indeed sequestered by the HMG box and basic regions acting together. In both cases, it is the concave DNA-binding face of the box that is occluded, and the tail remains largely disordered, as for HMGB1. We also establish a role for phosphorylation of the acidic tail in fine-tuning the degree of

auto-inhibition in HMG-D and ZmHMGB1; the sequence of the HMGB1 tail precludes such regulation.

## EXPERIMENTAL PROCEDURES

**Expression and Purification of <sup>15</sup>N-labeled HMG-box Proteins**—The plasmids encoding HMG-D  $\Delta$ 12 and the Ser  $\rightarrow$  Cys mutants HMG-D S103C and HMG-D S111C were generated from the wild-type plasmid (pET24a HMG-D112 (16)) by QuikChange mutagenesis (Agilent). Expression of all HMG-D constructs was carried out as described (30), using <sup>15</sup>NH<sub>4</sub>Cl as the sole nitrogen source, except that growth was continued for 12 h at 16 °C after induction.

Cell pellets were resuspended in 10 mM sodium phosphate (pH 6.0), 1 mM EDTA, 1 mM DTT, 0.5 mM PMSF, and 1 M NaCl with protease inhibitor mixture (5  $\mu$ g/ml aprotinin, pepstatin A, leupeptin; 0.78 mg/ml benzamidine; and 0.5 mM PMSF). Cells were disrupted by two passes through a French press (1000 p.s.i.), and the lysate was clarified by centrifugation. Ground (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> was added to 55% saturation (on ice, over 20 min). After a further 20 min on ice, the suspension was centrifuged, and the supernatant containing HMG-D was filtered (0.22  $\mu$ m, Millipore) and then dialyzed overnight against buffer A (10 mM sodium phosphate (pH 6.0), 1 mM EDTA, 1 mM DTT, 0.5 mM PMSF). The protein was then bound to a 5-ml HiLoad SP Sepharose HP column (GE Healthcare) and eluted using a linear gradient from buffer A to 70% buffer B (buffer A plus 1 M NaCl). Fractions containing HMG-D of sufficient purity (shown by SDS/18% polyacrylamide gel electrophoresis and Coomassie Blue staining) were pooled, concentrated, and exchanged into NMR buffer (10 mM sodium phosphate (pH 6.0), 1 mM EDTA) using a Vivaspin 2 concentrator (Sartorius; 5-kDa molecular mass cut-off). Final protein concentration was estimated from the A<sub>280</sub> using the extinction coefficient,  $\epsilon = 19,837 \text{ M}^{-1}\text{cm}^{-1}$ .

ZmHMGB1 was expressed from plasmid pT7 cm-HMGB1 (17). The sequence encoding ZmHMGB1  $\Delta$ 34 (residues Met<sup>1</sup>–Lys<sup>123</sup>) was amplified from this plasmid by PCR, digested with NdeI/HindIII, and inserted into NdeI/HindIII-digested vector pT7 cm. ZmHMGB1 and ZmHMGB1  $\Delta$ 34 were expressed in M9 medium supplemented with 8% <sup>15</sup>N Celtone medium (Cambridge Isotope Laboratories Inc.). The proteins were purified as described previously (17).

**Structural Model of the HMG Box of ZmHMGB1**—No high-resolution structure exists in the Protein Data Bank (PDB) for the HMG box of ZmHMGB1. However, given its high sequence homology to the B box of HMGB1 (the highest match of all HMG boxes in the PDB), a model of the ZmHMGB1 box (residues 35–112) could be constructed based on entry 1HME (8) using SWISS-MODEL (31–33). The resulting model was judged to be reliable based on the QMEAN4 score (34) of 0.85.

**NMR Spectroscopy**—NMR measurements were made on <sup>15</sup>N-labeled proteins (HMG-D full-length and  $\Delta$ 12, 1–2 mM; ZmHMGB1  $\Delta$ 34, 1.2 mM; ZmHMGB1, 0.04 mM) in 10% <sup>2</sup>H<sub>2</sub>O, 10 mM sodium phosphate (pH 7.0), 1 mM EDTA, 1 mM DTT, 0.5 mM PMSF (NMR buffer). Experiments were recorded at 30 °C (HMG-D) or 25 °C (ZmHMGB1) on Bruker DRX500, DRX600, or DRX800 spectrometers. Data were processed using the Azara suite of programs.<sup>5</sup> Assignments were made using CcpNmr Analysis v. 2.1 (35). Backbone assignments were derived from three-dimensional NOESY <sup>15</sup>N HSQC and TOCSY <sup>15</sup>N HSQC experiments (36). Chemical shift differences were calculated using  $\Delta\delta = [(\Delta\delta^H)^2 + (0.15 \times \Delta\delta^N)^2]^{1/2}$  (37). <sup>1</sup>H/<sup>15</sup>N heteronuclear NOE ratios ( $I_{\text{sat}}/I_{\text{unsat}}$ ) were obtained for <sup>15</sup>N-HMG-D at 500 MHz by deploying either 4 s of <sup>1</sup>H saturation using a 120° pulse train or a 4-s delay prior to the first <sup>15</sup>N pulse (38); errors were estimated from the standard deviation of the noise.

**Paramagnetic Relaxation Enhancement (PRE) Measurements**—<sup>15</sup>N-HMG-D S103C and S111C samples were incubated with 10 mM DTT at 25 °C for 15 min to ensure that the cysteine side chains were fully reduced. Spin label reagents (1-oxyl-2,2,5,5-tetramethyl-3-pyrrolidin-3-yl) methyl methanethiosulfonate (MTSL) (39) and [1-acetoxy-2,2,5,5-tetramethyl-3-pyrrolidin-3-yl] methyl methanethiosulfonate (ATSL) were added as described previously (29), and proteins exchanged into NMR buffer. Incorporation of the spin label into the protein was confirmed by electrospray ionization mass spectrometry (Dr. Len Packman, Protein and Nucleic Acid Chemistry Facility, Department of Biochemistry, University of Cambridge). PREs were extracted from the <sup>1</sup>H<sup>N</sup> transverse relaxation rate,  $\Gamma_2$ , via a two time-point measurement:  $\Gamma_2 = R_{2,\text{para}} - R_{2,\text{dia}} = [1/(T_b - T_a)] \times \ln[I_{\text{dia}}(T_b) I_{\text{para}}(T_a)/I_{\text{dia}}(T_a) I_{\text{para}}(T_b)]$ , where  $T_a$  and  $T_b$  were set to 12  $\mu$ s and 8 ms, respectively (40). Errors were estimated from the noise.

**Phosphorylation**—*In vitro* phosphorylation of HMG-D was carried out by incubating 0.25 mM protein (final volume of 500  $\mu$ l) at 25 °C with 5000 units of CK2 (New England Biolabs) and 1.25 mM ATP (5-fold molar excess) in 20 mM Tris-HCl (pH 7.5), 50 mM KCl, 10 mM MgCl<sub>2</sub>. Phosphorylation was monitored in

real time by <sup>15</sup>N HSQC, and by mass spectrometry (MALDI-in-source decay (MALDI-ISD) performed by Dr. Len Packman, as above). The sample was subsequently exchanged into NMR buffer for comparison with the non-phosphorylated protein.

*In vitro* phosphorylation of <sup>15</sup>N-ZmHMGB1 was achieved by incubation of 40  $\mu$ M protein in a total volume of 600  $\mu$ l at 25 °C overnight with 0.4  $\mu$ M recombinant *Z. mays* CK2 (expressed and purified as described (23)) and 400  $\mu$ M ATP in CK2 buffer (25 mM Tris-HCl (pH 8.5), 10 mM MgCl<sub>2</sub>, 1 mM DTT) (24). The extent of phosphorylation was checked by mass spectrometry (as above) and NMR spectroscopy (after exchange into NMR buffer).

## RESULTS

Comparison of the <sup>15</sup>N HSQC spectra of full-length and tail-less HMG-D and ZmHMGB1 should reveal any intra-molecular contacts of the tail with the remainder of the protein, through chemical shift perturbations, as in the case of vertebrate HMGB1 (28, 29). The contacts in HMG-D were also explored by PRE measurements, which provide a more direct measurement of distance and are less sensitive to conformational changes. The effect of tail phosphorylation was also investigated for both proteins.

**Chemical Shift Perturbation Mapping**—<sup>15</sup>N HSQC spectra of HMG-D are typical of HMG-box-containing proteins (28), having a combination of well dispersed peaks attributable to the HMG box and sharper, overlapping peaks in the central region, corresponding to the disordered basic linker and acidic tail (Fig. 2a). There is also evidence of conformational exchange on a  $\mu$ s–ms timescale for some peaks that are broad and weak. No evidence of intermolecular interactions was detected (<sup>15</sup>N HSQC spectra over a concentration range of 0.1–1 mM overlay exactly (not shown)). Chemical shift assignments were obtained using three-dimensional NOESY- and TOCSY-HSQC experiments recorded on <sup>15</sup>N-labeled HMG-D and HMG-D  $\Delta$ 12. Of the 107 non-proline backbone <sup>1</sup>H/<sup>15</sup>N pairs in full-length HMG-D and 95 in HMG-D  $\Delta$ 12, 91 and 78 pairs, respectively, could be unambiguously assigned.

Chemical shift differences between full-length HMG-D and the tail-deleted protein are small ( $\Delta\delta < 0.2$  ppm) but widespread (Fig. 2, a and b). The mean shift difference (excluding the C-terminal residue, Lys<sup>100</sup>, adjacent to the deletion) is  $0.028 \pm 0.023$  ppm (Fig. 2). This is much smaller than the value of  $0.045 \pm 0.062$  ppm for HMGB1 (28). Nevertheless, as in the case of HMGB1, the residues that show the greatest shifts map to helices I and II, and also to a patch at the C-terminal end of helix III (in particular residue Lys<sup>67</sup>) and the adjacent portion of the  $\beta$ -strand; most of these residues lie on the concave DNA-binding face of the protein (Fig. 2b).

For ZmHMGB1 (Fig. 3), peak overlap and broadening due to intermediate exchange on the chemical shift timescale precludes a complete assignment. This is most problematic in the disordered tail and linker regions where inter-residue NOE information is both sparse and ambiguous. However, a reasonable distribution of representative assignments was obtained for residues 1–123 encompassing the N-terminal tail, HMG box, and basic linker; 72 and 78, respectively, out of a possible 115 non-proline backbone <sup>1</sup>H/<sup>15</sup>N pairs could be unambigu-

<sup>5</sup> Azara, v. 2.8, copyright © 1993–2014 Wayne Boucher and Department of Biochemistry, University of Cambridge, UK.

## Role of the Acidic Tail of Single HMG-box Proteins

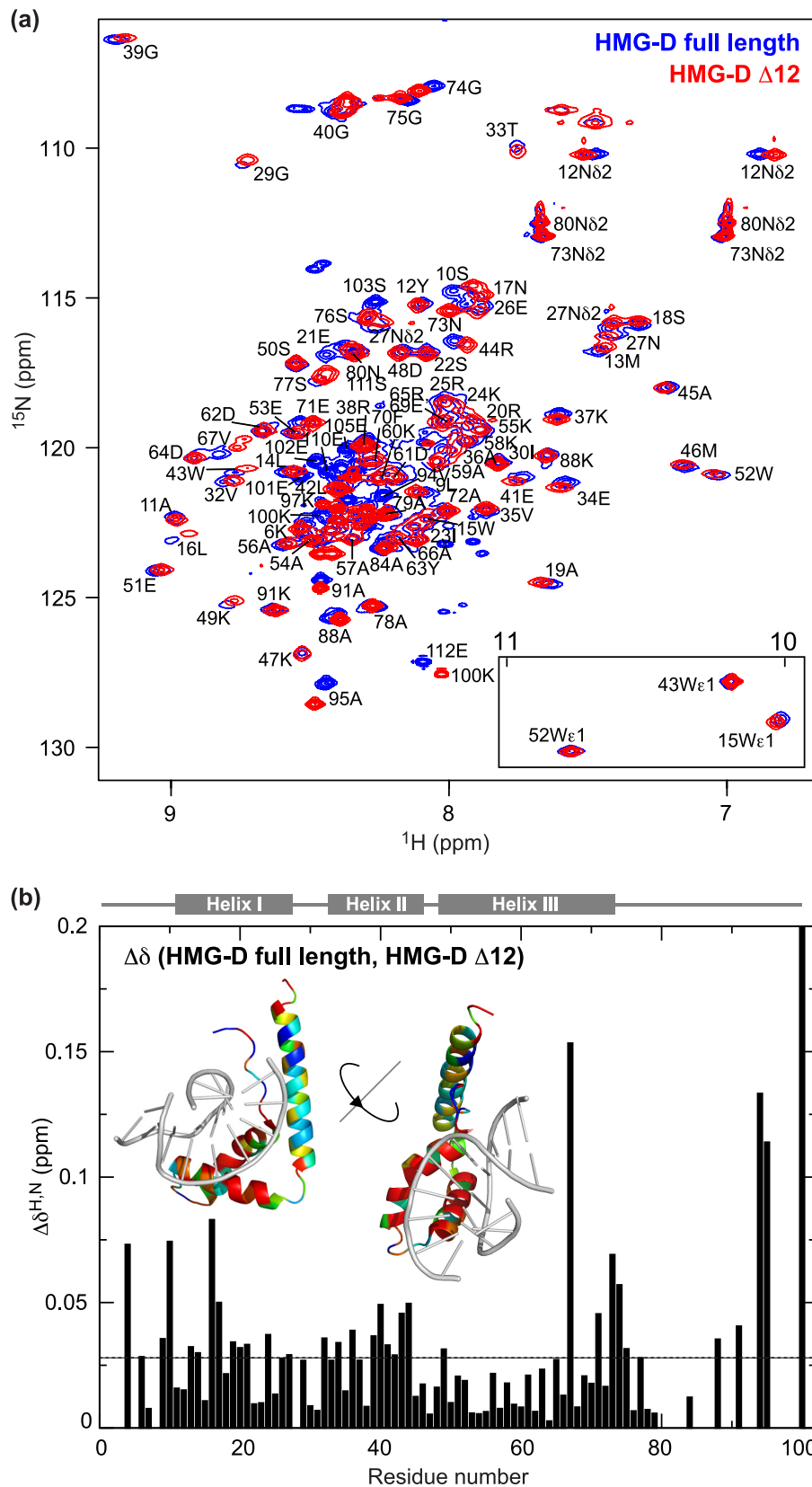


FIGURE 2. **NMR spectroscopy of HMG-D and HMG-D lacking the acidic tail (HMGD  $\Delta 12$ ).** *a*,  $^{15}\text{N}$  HSQC spectra. *b*, chemical shift differences ( $\Delta\delta = [(\Delta\delta^H)^2 + (0.15 \times \Delta\delta^N)^2]^{1/2}$  (37)); mean  $\Delta\delta$  0.028 ppm (dotted line). Inset ribbon structures are based on 1QRV.  $\text{Log}_{10}(\Delta\delta + 1)$  was converted to a rainbow color ramp (blue = least shifted, red = most shifted residues); DNA is shown (gray) to demonstrate the similarity between the acidic-tail-binding and DNA-binding surfaces.

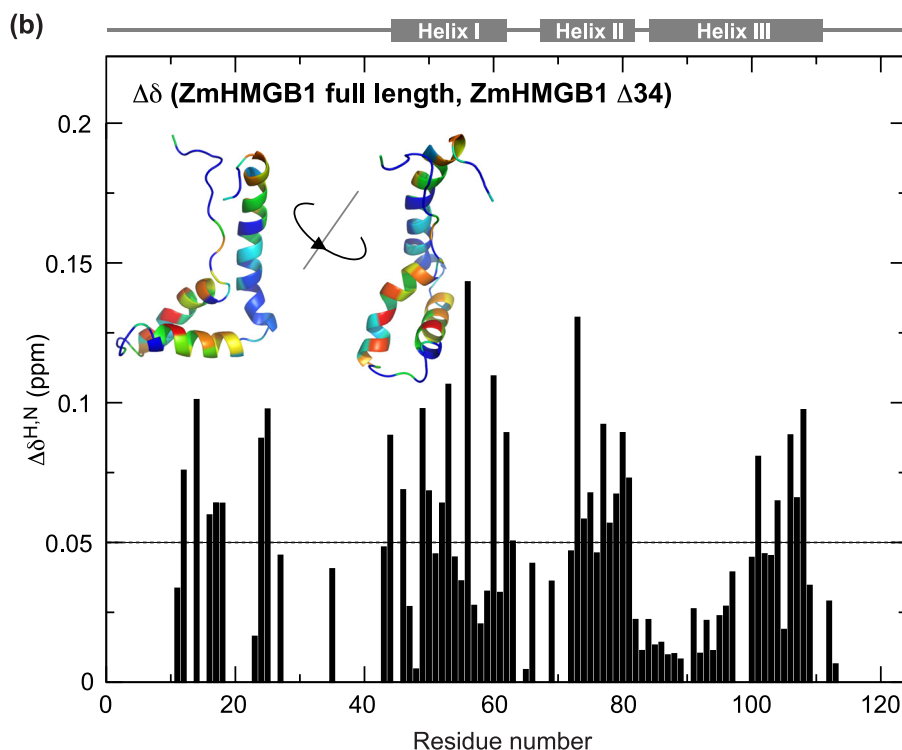
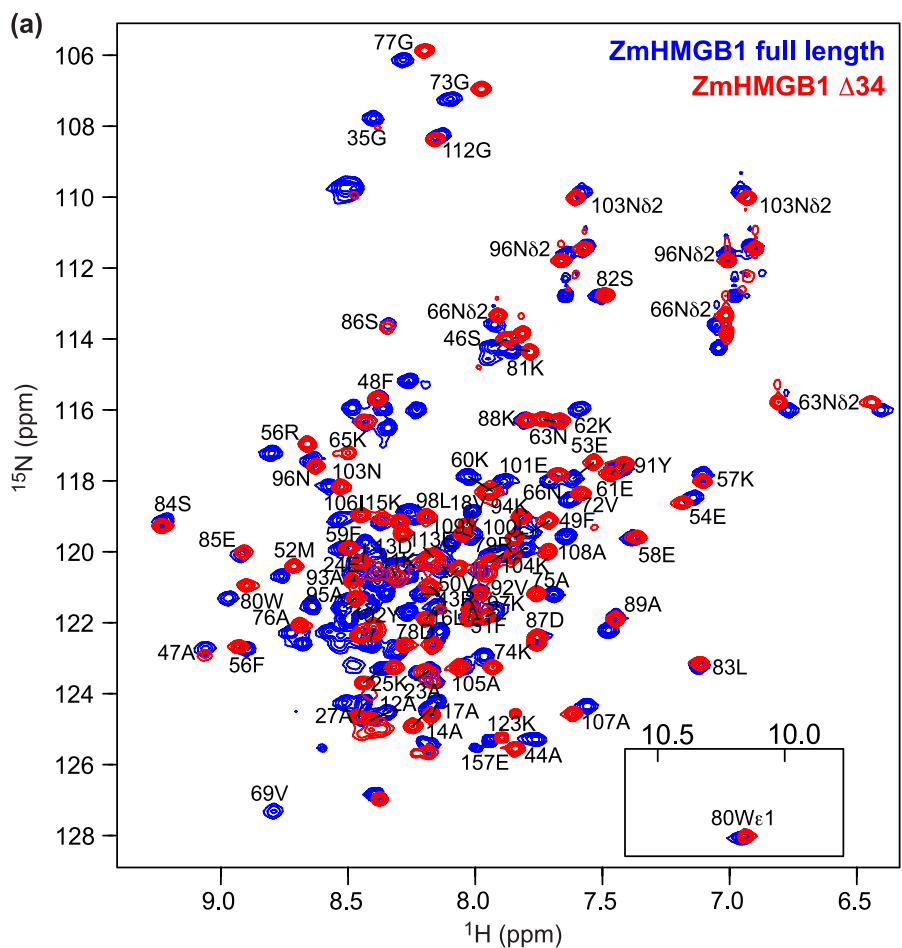


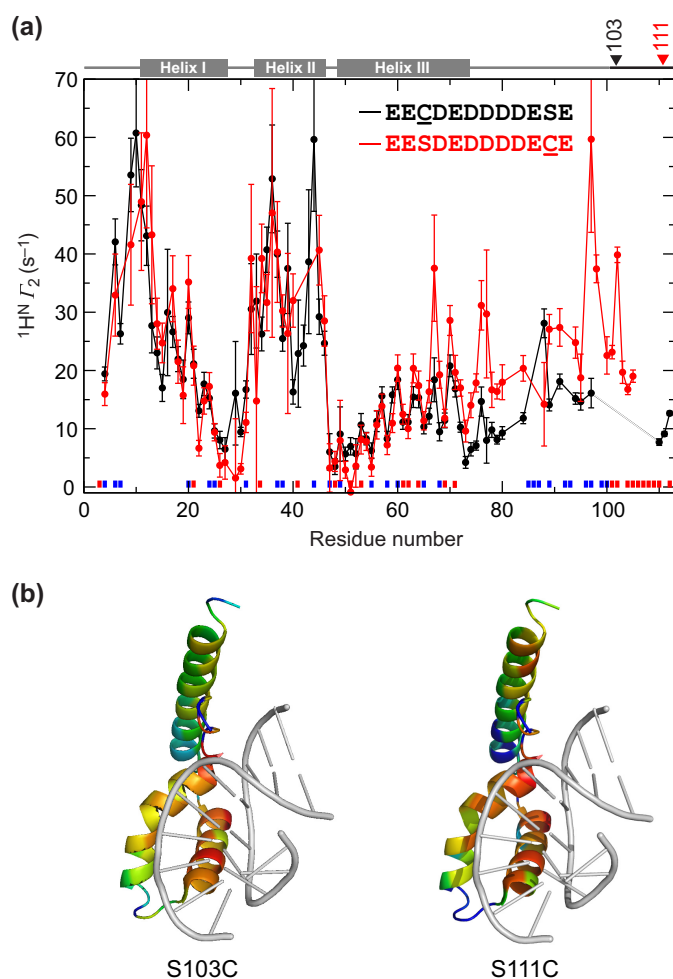
FIGURE 3. **NMR spectroscopy of ZmHMGB1 and ZmHMGB1 lacking the acidic tail (ZmHMGB1  $\Delta$ 34).** *a*,  $^{15}\text{N}$  HSQC spectra. *b*, chemical shift differences ( $\Delta\delta = [(\Delta\delta^{\text{H}})^2 + (0.15 \times \Delta\delta^{\text{N}})^2]^{1/2}$  (37)); mean  $\Delta\delta$  0.050 ppm (dotted line). Inset ribbon structures are based on the homology model (see "Experimental Procedures").  $\text{Log}_{10}(\Delta\delta + 1)$  was converted to a rainbow color ramp (blue = least shifted, red = most shifted residues).

## Role of the Acidic Tail of Single HMG-box Proteins

ously assigned for ZmHMGB1 and ZmHMGB1  $\Delta$ 34. Only four out of a possible 34 backbone  $^1\text{H}/^{15}\text{N}$  pairs could be assigned unambiguously in the acidic C-terminal tail, due to the extensive overlap of the 25 Asp and Glu residues in this disordered and sequence-repetitive region. Chemical shift differences between the full-length and tail-less proteins (Fig. 3, *a* and *b*; mean  $\Delta\delta$   $0.050 \pm 0.033$  ppm) are more pronounced than for HMG-D and closer to the values seen in HMGB1, and again they map primarily to the concave DNA-binding face of the HMG box (Fig. 3*b*). Residues that could be assigned in the basic N-terminal tail shift to a similar extent to those in the HMG box and the shifts are well distributed over the length of the tail (Fig. 3*b*). It may be significant that the relative mean changes in chemical shift for HMG-D, ZmHMGB1, and HMGB1 (0.028, 0.050, and 0.045 ppm, respectively) increase with increasing tail length (12, 34, and 30 residues, respectively), presumably reflecting increasing affinity of the tails.

**Paramagnetic Relaxation Enhancement**—To map contacts made by the (short) acidic tail of HMG-D using PREs, two serine-to-cysteine mutants (HMG-D S103C and S111C) were generated to permit attachment of paramagnetic probes near the distal and proximal ends of the tail (residues 101–112). The proteins were labeled with either the nitroxide-containing MTSL or its diamagnetic analog, ATSL, as carried out previously with HMGB1 (29). PREs were then quantified as the amide proton transverse paramagnetic relaxation rate enhancement ( $^1\text{H}^N \Gamma_2 = R_{2,\text{para}} - R_{2,\text{dia}}$ ) (40) for each mutant (Fig. 4*a*). Pronounced reductions in  $\Gamma_2$ , indicative of proximity to the paramagnetic tag, are seen for a similar set of residues as in the chemical shift mapping (Fig. 4*b*), the main difference being a less dramatic effect seen by PRE at the C terminus of helix III, in particular for Lys<sup>67</sup>. One explanation might be that the changes seen by chemical shift mapping are due in part to a conformational change away from the binding site rather than a direct interaction with the acidic tail (similar chemical shift changes are seen for the B box of HMGB1 where C-terminal truncation destabilizes the last turn of the helix (41)). Another might be that the acidic tail competes for the basic linker with the acidic patch at the C terminus of helix III (Fig. 1; not present in HMGB1 or ZmHMGB1), and thus the changes seen actually correspond to detachment of the basic linker. The tail appears to make contact with both the basic linker and the HMG box.

**Phosphorylation**—Two constitutive HMG-D phosphorylation sites have been detected *in vivo* and inferred to be Ser<sup>103</sup> and Ser<sup>111</sup> due to their location in CK2 recognition motifs (22). To investigate the effect of phosphorylation on HMG-D, and any structural impact,  $^{15}\text{N}$ -labeled protein was phosphorylated using CK2. The reaction was followed in real time by NMR (Fig. 5). Phosphorylation was inferred to occur at both Ser<sup>103</sup> and Ser<sup>111</sup> as the intensity of peaks corresponding to these residues dropped over time, and corresponding peaks appeared in markedly different positions (Fig. 5*a*). This was shown directly by mass spectrometry (mass increases of 80 Da at serines 103 and 111 in sequential fragment analysis by MALDI-in-source decay (not shown)). Phosphorylation was essentially complete after 3 h, as judged by the absence of further spectral changes. Ser<sup>103</sup> was phosphorylated more rapidly than Ser<sup>111</sup> (Fig. 5*b*), probably due to its location in a complete CK2 recognition motif

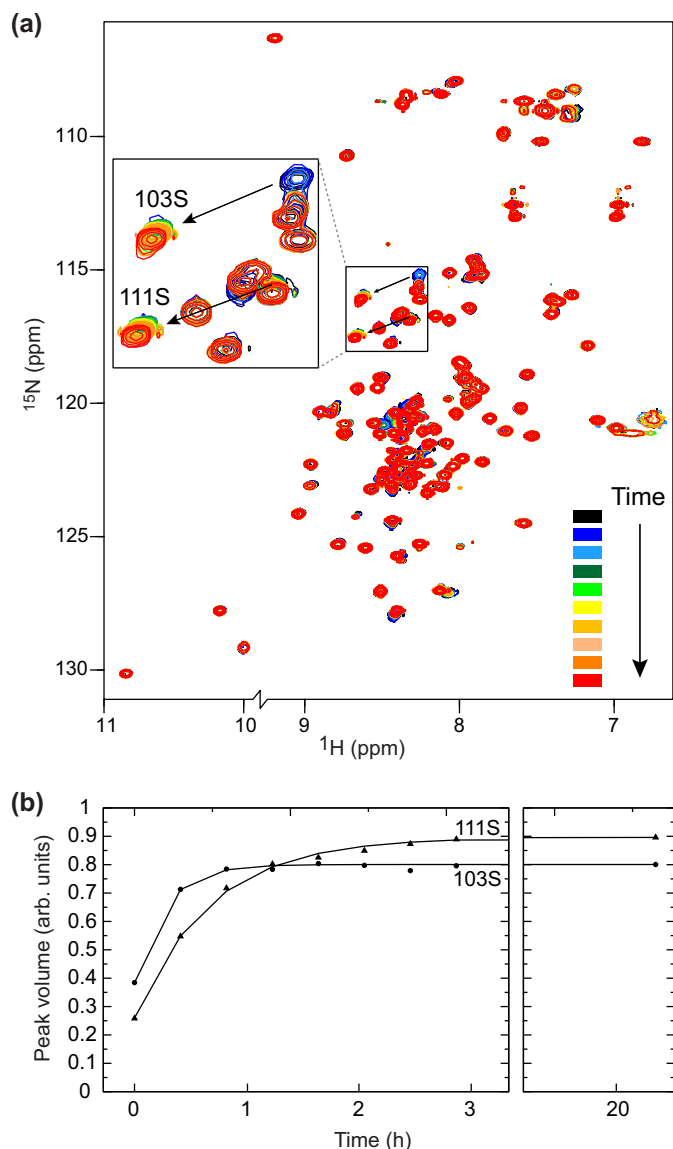


**FIGURE 4. Paramagnetic relaxation enhancement study of HMG-D Ser  $\rightarrow$  Cys mutants.** *a*, PRE values measured as  $^1\text{H}^N \Gamma_2$ ; HMG-D S103C is in black, and S111C is in red. Red and blue blocks above the x axis indicate the positions of acidic and basic residues, respectively. *b*, ribbon structures are based on 1QRV.  $\text{Log}_{10}(\Gamma_2)$  was converted to a rainbow color ramp (blue = lowest, red = highest) encompassing both the S103C and the S111C datasets; DNA is shown (gray) to demonstrate the similarity between the acidic-tail-binding and DNA-binding surfaces.

SXX(D/E) ( $^{103}\text{SDED}$ ) as compared with  $^{111}\text{SE}$  at the C terminus (42).

Peaks from residues adjacent to the phosphorylation sites also shifted slightly, along with residues throughout the HMG box (mean  $\Delta\delta$   $0.026 \pm 0.018$  ppm, similar in magnitude to  $0.028 \pm 0.023$  ppm seen for tail deletion; Fig. 6*a*). Strikingly, the trajectory of the shifts was almost uniformly in the opposite direction to that following tail deletion (Fig. 6*b*), suggesting that the effect of phosphorylation is to enhance tail binding.

ZmHMGB1 is phosphorylated at 4 residues *in vivo* (cultured maize cells and immature maize kernels). Three sites were mapped to the acidic tail and could be phosphorylated by CK2 *in vitro*; Ser<sup>149</sup> was fully phosphorylated, whereas Ser<sup>133</sup> and Ser<sup>136</sup> were phosphorylated less efficiently (75% of the protein was phosphorylated at both sites (23, 24)). We phosphorylated  $^{15}\text{N}$ -ZmHMGB1 using *Z. mays* CK2, and although the low sample concentration made it impractical to follow the reaction in real time, comparison of the fully phosphorylated protein (denoted Zm-HMGB1-P) with the unmodified protein (Fig. 6*c*)



**FIGURE 5. Real-time phosphorylation of  $^{15}\text{N}$ -labeled HMG-D monitored by  $^{15}\text{N}$  HSQC NMR spectroscopy.** *a*, spectra are colored black through to red to indicate the increasing time. Peaks arising from Ser<sup>103</sup> and Ser<sup>111</sup> undergo a significant chemical shift change shown by the arrow, indicative of phosphorylation. The intensity of the original peak falls with time, whereas the intensity of the corresponding peak in the new position grows. *b*, quantification of the intensity of the new peaks seen in *a*. arb. units, arbitrary units.

showed a major shift in the position of the peak arising from Ser<sup>149</sup>; this was of a similar magnitude and along a similar trajectory to that of the serine residues in HMG-D. An unassigned serine in ZmHMGB1 also shifted, most probably one of the serines (Ser<sup>133</sup> and Ser<sup>136</sup>) previously observed to be less efficiently phosphorylated (23), or possibly both if they have overlapping chemical shifts. As in the case of HMG-D, adjacent residues were shifted slightly, but there were also widespread small changes over much of the N-terminal tail and HMG box (Fig. 6c), generally in the opposite direction to that caused by tail deletion (not shown), again indicating that phosphorylation results in an enhancement of tail binding. The magnitude of the changes (mean  $\Delta\delta$   $0.014 \pm 0.010$  ppm) was less than for tail deletion ( $0.050 \pm 0.033$  ppm).

Phosphorylation has less of an effect on tail interactions in ZmHMGB1, as judged by chemical shifts, than in HMG-D (mean  $\Delta\delta$  0.026), possibly because the density of phosphorylation sites is somewhat lower in the longer ZmHMGB1 tail (three sites in 34 residues for ZmHMGB1 versus two sites in 12 residues for HMG-D).

**Backbone Dynamics**—Both the basic (AK-rich) regions flanking the HMG box(es) and the acidic tails are intrinsically disordered in HMG-box proteins. Interestingly, on the evidence currently available, both appear to remain largely unstructured on binding to DNA or other proteins; the HMG-D basic linker binds to distorted (disulfide-cross-linked) DNA without induction of any regular secondary structure (20), and no induced structure was detected for the acidic tail of HMGB1 on binding to the basic tails of histones H1 (43) or H3 (44).

Because chemical shift mapping suggests that phosphorylation enhances tail binding, we performed  $\{^1\text{H}\}^{15}\text{N}$  heteronuclear NOE experiments on unmodified  $^{15}\text{N}$ -HMG-D and the fully phosphorylated protein (denoted  $^{15}\text{N}$ -HMG-D-P) to establish whether phosphorylation was accompanied by a change in fast ns-ps motions in the disordered regions of the protein (Fig. 7). (This was not practical for ZmHMGB1 due to lack of available material.)  $^{15}\text{N}$ -HMG-D showed relatively high heteronuclear NOE values for the regions of known secondary structure, with a small dip corresponding to the loop region between helices I and II. Unusually, the most flexible region with large negative values was the AG-rich region preceding the basic linker, suggesting that it may be “looped out” of the assembled whole. The N- and C-terminal regions (which have  $\{^1\text{H}\}^{15}\text{N}$  NOE values around zero), although more flexible than the box, were less dynamic than the AG-rich loop ( $\{^1\text{H}\}^{15}\text{N}$  NOE  $\sim -1$ ), consistent with their involvement in intramolecular interactions while retaining a degree of disorder.

Phosphorylation resulted in local decreases in the extent of motions faster than overall tumbling. The effect was small and non-uniform for the regions of established secondary structure, but significant and consistent in the disordered regions. This is very likely to be due to an increase in the affinity of the basic regions for the phosphorylated tail, which carries two extra negative charges. Similar changes in the opposite direction were seen previously for HMGB1 upon the addition of salt, when the effect of charge screening by  $\text{Na}^+$  and  $\text{Cl}^-$  ions was to weaken acidic tail interactions (29). An additional experiment was therefore performed on unphosphorylated HMG-D after the addition of 200 mM NaCl. Judging by the effect on fast motions revealed by the  $\{^1\text{H}\}^{15}\text{N}$  NOE, salt appeared to reduce the tail affinity by an amount similar in magnitude to the removal of the two phosphate groups.

## DISCUSSION

Despite the differences between the abundant HMG-box proteins from various species, the mode of auto-inhibition is conserved; the acidic tails in all cases map to an analogous interface with the HMG box(es) and basic tail/linker regions (Fig. 8). The overall effect on the structure is necessarily less dramatic for the single-box proteins than for HMGB1, where two tandem boxes are drawn together by the acidic tail to form a relatively compact, dynamic, assembly. However, in the single-box

## Role of the Acidic Tail of Single HMG-box Proteins

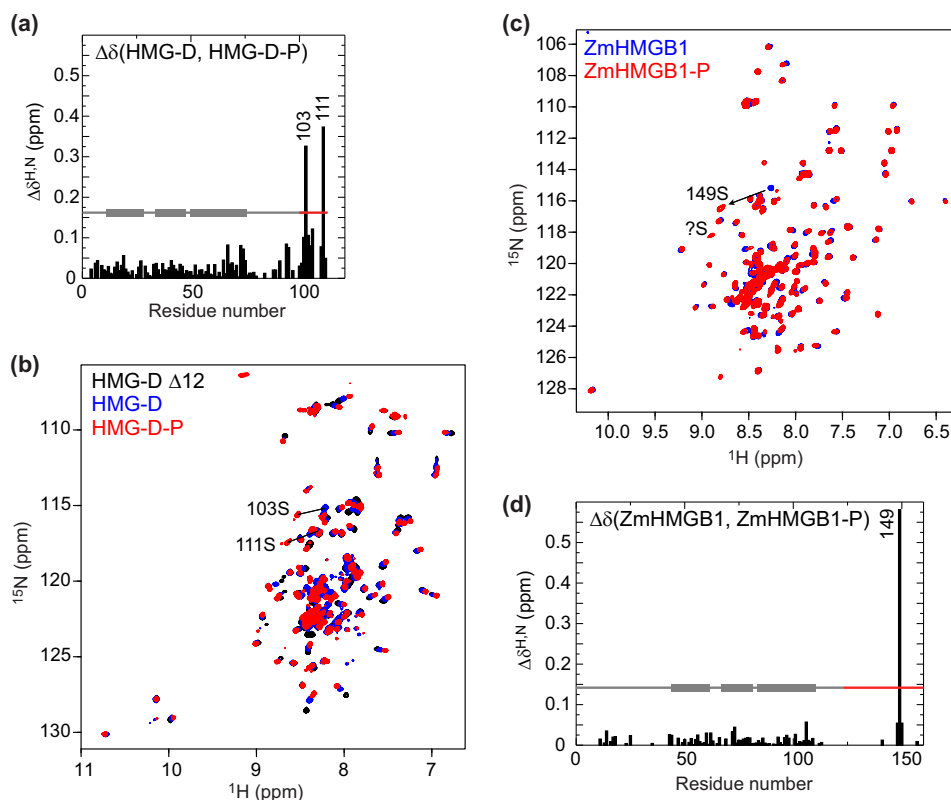


FIGURE 6.  $^{15}\text{N}$  HSQC spectra of single-box HMG proteins in the absence and presence of phosphorylation. *a*, chemical shift differences ( $\Delta\delta = [(\Delta\delta^{\text{H}})^2 + (0.15 \times \Delta\delta^{\text{N}})^2]^{1/2}$  (37)) between HMG-D and the phosphorylated form, HMG-D-P; Ser<sup>103</sup> and Ser<sup>111</sup> are phosphorylated (Fig. 5). *b*,  $^{15}\text{N}$  HSQC spectra of HMG-D (blue), HMG-D-P (red), and HMG-D  $\Delta$ 12 (black) for comparison of shift trajectories. *c*, ZmHMGB1 (blue) and the phosphorylated form, ZmHMGB1-P (red); Ser<sup>149</sup> undergoes a significant chemical shift change indicative of phosphorylation, as does another peak that cannot be assigned due to peak overlap in the spectrum of the unphosphorylated protein. *d*, chemical shift differences as in *a*.

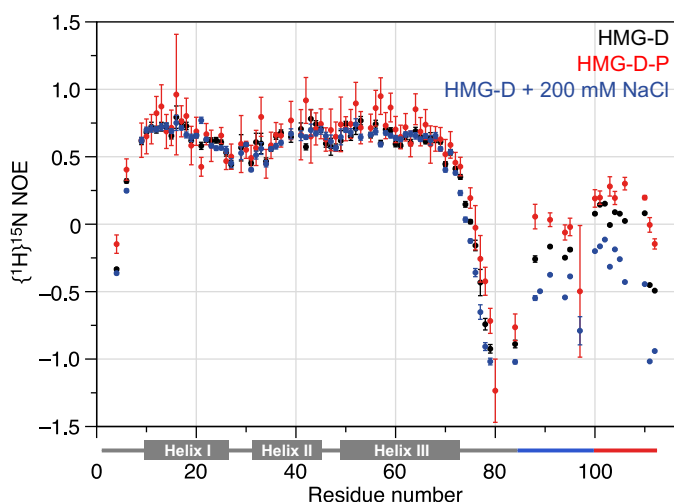


FIGURE 7. Backbone dynamics of HMG-D (black), HMG-D-P (red) in low salt buffer, and HMG-D in 200 mM NaCl (blue).  $\{^1\text{H}\}^{15}\text{N}$  Heteronuclear NOE values show a reduction in the extent of flexible motions in the N- and C-terminal tail regions upon phosphorylation, whereas the addition of salt results in an increase. (Note: Some error bars for residues in the flexible regions are not visible due to the high signal-to-noise ratio.)

proteins, tail phosphorylation (not possible for vertebrate HMGB1) potentially adds an extra level of regulatory control.

The effect of phosphorylation on the conformation of the short acidic tails of the *Drosophila* and maize HMG-box proteins appears to involve a tightening of the existing interaction with the boxes and basic tail/linker regions. This may have con-

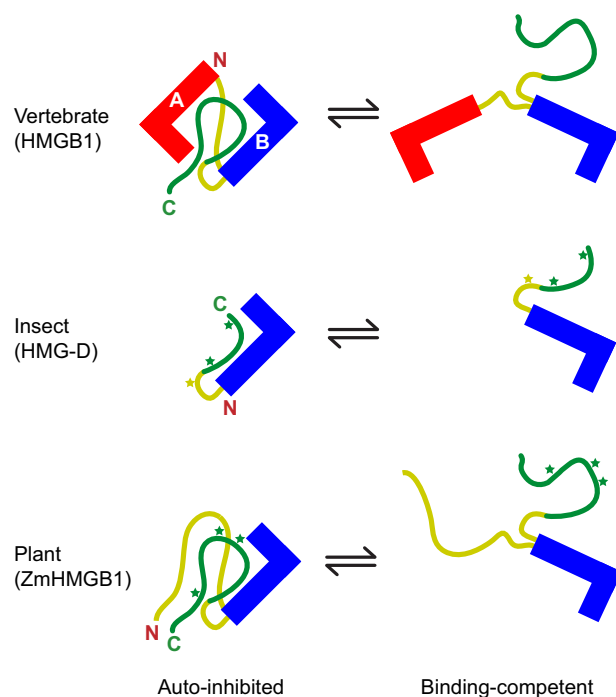


FIGURE 8. Schematic showing the closed, auto-inhibited versus open, binding-competent forms of the HMG proteins studied to date. A-type boxes are shown in red, B-type are shown in blue, basic linker regions are shown in yellow, and acidic tail regions are shown in green. Stars indicate sites phosphorylated by CK2 (green) or PKC (yellow).



sequences for protein function (e.g. phosphorylation disrupts the maize HMGB1-Dof2 complex, abrogating HMG-facilitated DNA binding of Dof2 (45, 46)). The local flexibility of the acidic tails is reduced on binding to the boxes and linkers/basic tail, and reduced further upon phosphorylation, but the tail retains a degree of disorder (Fig. 7), as in the case of the interaction of (phosphorylated) ubiquitin ligase with the intrinsically disordered N-terminal domain of its inhibitor Sic1 (47) and other “fuzzy complexes” (48).

For HMGB proteins, there is no detectable increase in canonical secondary structure in the free proteins on phosphorylation. However, there might be changes in structure on complex formation that are influenced by phosphorylation. For example, the basic disordered histone H1 C-terminal domain becomes structured on binding to DNA, and phosphorylation results in a decrease in the DNA-dependent  $\alpha$ -helical content and a corresponding increase in  $\beta$ -structures (49).

For HMG-D, which has been well studied, the tail promotes the transient unwrapping of nucleosomal DNA (50), presumably in part through binding to the N-terminal region of H3 (44, 51); phosphorylation would be expected to influence this unwrapping and alter nucleosome stability, and consequently affect chromatin higher-order structure. HMG-D substitutes for H1 in mitotic chromosomes in the earliest stages of *Drosophila* embryogenesis (52, 53). We speculate that HMG-D phosphorylation/dephosphorylation might be key to controlling compaction/decompaction of the chromatin fiber in *Drosophila* embryos (essentially the entire population of HMG-D and its homologue HMG-Z is phosphorylated in 0–18-h embryos (22)). There might be parallels with phosphorylation of the H1 C-terminal domain in mitotic chromosomes in higher organisms (54), in marked contrast to the unphosphorylated state of linker histones in other forms of highly condensed (amitotic) chromatin of, for example, chicken erythrocyte nuclei (55, 56) and sea urchin sperm (57, 58). Exactly how phosphorylation relates to mitotic condensation remains to be established. We suggest that phosphorylation of the acidic tail of single HMG-box proteins might modulate protein function by shifting the equilibrium toward the tail-bound auto-inhibited state (29). Moreover, phosphorylation may also directly affect the structure and interaction of HMGB proteins with partners, by creating or disrupting binding motifs, providing an additional level of regulation.

*Acknowledgment*—We thank Fred Northrop for technical assistance.

## REFERENCES

- 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
- Bustin, M. (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol. Cell Biol.* **19**, 5237–5246
- Thomas, J. O., and Travers, A. A. (2001) HMG1 and 2, and related “architectural” DNA-binding proteins. *Trends Biochem. Sci.* **26**, 167–174
- Agresti, A., and Bianchi, M. (2003) HMGB proteins and gene expression. *Curr. Op. Genet. Dev.* **13**, 170–178
- Travers, A. A. (2003) Priming the nucleosome: a role for HMGB proteins. *EMBO Rep.* **4**, 131–136

- Štros, M. (2010) HMGB proteins: interactions with DNA and chromatin. *Biochim. Biophys. Acta* **1799**, 101–113
- 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
- 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 HMG-1. *EMBO J.* **12**, 1311–1319
- Murphy, F. V., 4th, Sweet, R. M., and Churchill, M. E. (1999) The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition. *EMBO J.* **18**, 6610–6618
- DeLano, W. L. (2010) *The PyMOL Molecular Graphics System*, version 0.99rc6, Schrödinger, LLC, New York
- Štros, M., Launholt, D., and Grasser, K. D. (2007) The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell. Mol. Life Sci.* **64**, 2590–2606
- Teo, S. H., Grasser, K. D., and Thomas, J. O. (1995) Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *Eur. J. Biochem.* **230**, 943–950
- Grasser, K. D., Teo, S. H., Lee, K. B., Broadhurst, R. W. B., Rees, C., Hardman, C. H., and Thomas, J. O. (1998) DNA-binding properties of the tandem HMG boxes of high-mobility-group protein 1 (HMG1). *Eur. J. Biochem.* **253**, 787–795
- Štros, M. (1998) DNA bending by the chromosomal protein HMG1 and its high mobility group box domains. *J. Biol. Chem.* **273**, 10355–10361
- Wiśniewski, J. R., and Schulze, E. (1994) High affinity interaction of dip-teran high mobility group (HMG) proteins 1 with DNA is modulated by COOH-terminal regions flanking the HMG box domain. *J. Biol. Chem.* **269**, 10713–10719
- Payet, D., and Travers, A. (1997) The acidic tail of the high mobility group protein HMG-D modulates the structural selectivity of DNA binding. *J. Mol. Biol.* **266**, 66–75
- Ritt, C., Grimm, R., Fernandez, S., Alonso, J. C., and Grasser, K. D. (1998) Basic and acidic regions flanking the HMG domain of maize HMGa modulate the interactions with DNA and the self-association of the protein. *Biochemistry* **37**, 2673–2681
- Lee, K. B., and Thomas, J. O. (2000) The effect of the acidic tail on the DNA-binding properties of the HMG1, 2 class of proteins: insights from tail switching and tail removal. *J. Mol. Biol.* **304**, 135–149
- Grasser, K. D. (1998) HMG1 and HU proteins: architectural elements in plant chromatin. *Trends Plant Sci.* **3**, 260–265
- Dow, L. K., Jones, D. N., Wolfe, S. A., Verdine, G. L., and Churchill, M. E. (2000) Structural studies of the high mobility group globular domain and basic tail of HMG-D bound to disulfide cross-linked DNA. *Biochemistry* **39**, 9725–9736
- Pearson, R. B., and Kemp, B. E. (1991) Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. *Methods Enzymol.* **200**, 62–81
- Wiśniewski, J., Szwczuk, Z., Petry, I., Schwanbeck, R., and Renner, U. (1999) Constitutive phosphorylation of the acidic tails of the high mobility group 1 proteins by casein kinase II alters their conformation, stability, and DNA binding specificity. *J. Biol. Chem.* **274**, 20116–20222
- Stemmer, C., Schwander, A., Bauw, G., Fojan, P., and Grasser, K. D. (2002) Protein kinase CK2 differentially phosphorylates maize chromosomal high mobility group B (HMGB) proteins modulating their stability and DNA interactions. *J. Biol. Chem.* **277**, 1092–1098
- Stemmer, C., Leeming, D. J., Franssen, L., Grimm, R., and Grasser, K. D. (2003) Phosphorylation of maize and *Arabidopsis* HMGB proteins by protein kinase CK2 $\alpha$ . *Biochemistry* **42**, 3503–3508
- Thomsen, M. S., Franssen, L., Launholt, D., Fojan, P., and Grasser, K. D. (2004) Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein. *Biochemistry* **43**, 8029–8037
- Pedersen, D. S., Merkle, T., Marktl, B., Lildballe, D. L., Antosch, M., Bergmann, T., Tönsing, K., Anselmetti, D., and Grasser, K. D. (2010) Nucleocytoplasmic distribution of the *Arabidopsis* chromatin-associated

## Role of the Acidic Tail of Single HMG-box Proteins

- HMGB2/3 and HMGB4 proteins. *Plant Physiol.* **154**, 1831–1841
27. Wiśniewski, J. R., Schulze, E., and Sapetto, B. (1994) DNA binding and nuclear translocation of insect high-mobility-group-protein-1 (HMG1) proteins are inhibited by phosphorylation. *Eur. J. Biochem.* **225**, 687–693
  28. Watson, M., Stott, K., and Thomas, J. O. (2007) Mapping intramolecular interactions between domains in HMGB1 using a tail-truncation approach. *J. Mol. Biol.* **374**, 1286–1297
  29. Stott, K., Watson, M., Howe, F. S., Grossmann, J. G., and Thomas, J. O. (2010) Tail-mediated collapse of HMGB1 is dynamic and occurs via differential binding of the acidic tail to the A and B domains. *J. Mol. Biol.* **403**, 706–722
  30. Jones, D. N., Searles, M. A., Shaw, G. L., Churchill, M. E., Ner, S. S., Keeler, J., Travers, A. A., and Neuhaus, D. (1994) The solution structure and dynamics of the DNA-binding domain of HMG-D from *Drosophila melanogaster*. *Structure* **2**, 609–627
  31. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **37**, D387–D392
  32. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201
  33. Peitsch, M. C. (1995) Protein modeling by E-mail. *Nat. Biotechnol.* **13**, 658–660
  34. Benkert, P., Biasini, M., and Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**, 343–350
  35. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696
  36. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, pp. 447–457, Academic Press, San Diego, CA
  37. Zuiderweg, E. R. (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* **41**, 1–7
  38. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by  $^{15}\text{N}$  NMR relaxation. *Biochemistry* **33**, 5984–6003
  39. Berliner, L. J., Grunwald, J., Hankovszky, H. O., and Hideg, K. (1982) A novel reversible thiol-specific spin label: papain active site labeling and inhibition. *Anal. Biochem.* **119**, 450–455
  40. Iwahara, J., Tang, C., and Marius Clore, G. (2007) Practical aspects of  $^1\text{H}$  transverse paramagnetic relaxation enhancement measurements on macromolecules. *J. Magn. Reson.* **184**, 185–195
  41. Stott, K., Tang, G. S., Lee, K. B., and Thomas, J. O. (2006) Structure of a complex of tandem HMG boxes and DNA. *J. Mol. Biol.* **360**, 90–104
  42. Meggio, F., and Pinna, L. A. (2003) One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **17**, 349–368
  43. Cato, L., Stott, K., Watson, M., and Thomas, J. O. (2008) The interaction of HMGB1 and linker histones occurs through their acidic and basic tails. *J. Mol. Biol.* **384**, 1262–1272
  44. Watson, M., Stott, K., Fischl, H., Cato, L., and Thomas, J. O. (2014) Characterization of the interaction between HMGB1 and H3: a possible means of positioning HMGB1 in chromatin. *Nucleic Acids Res.* **42**, 848–859
  45. Krohn, N. M., Yanagisawa, S., and Grasser, K. D. (2002) Specificity of the stimulatory interaction between chromosomal HMG proteins and the transcription factor Dof2 and its negative regulation by protein kinase CK2-mediated phosphorylation. *J. Biol. Chem.* **277**, 32438–32444
  46. Grasser, M., Christensen, J. M., Peterhänsel, C., and Grasser, K. D. (2007) Basic and acidic regions flanking the HMG-box domain of maize HMGB1 and HMGB5 modulate the stimulatory effect on the DNA binding of transcription factor Dof2. *Biochemistry* **46**, 6375–6382
  47. Mittag, T., Orlicky, S., Choy, W.-Y., Tang, X., Lin, H., Sicheri, F., Kay, L. E., Tyers, M., and Forman-Kay, J. D. (2008) Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17772–17777
  48. Tompa, P., and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **33**, 2–8
  49. Roque, A., Ponte, I., Arrondo, J. L., and Suau, P. (2008) Phosphorylation of the carboxy-terminal domain of histone H1: effects on secondary structure and DNA condensation. *Nucleic Acids Res.* **36**, 4719–4726
  50. Ragab, A., and Travers, A. (2003) HMG-D and histone H1 alter the local accessibility of nucleosomal DNA. *Nucleic Acids Res.* **31**, 7083–7089
  51. Thomas, J. O., and Stott, K. (2012) H1 and HMGB1: modulators of chromatin structure. *Biochem. Soc. Trans.* **40**, 341–346
  52. Ner, S. S., and Travers, A. A. (1994) HMG-D, the *Drosophila melanogaster* homologue of HMG 1 protein, is associated with early embryonic chromatin in the absence of histone H1. *EMBO J.* **13**, 1817–1822
  53. Ner, S. S., Blank, T., Pérez-Paralle, M. L., Grigliatti, T. A., Becker, P. B., and Travers, A. A. (2001) HMG-D and histone H1 interplay during chromatin assembly and early embryogenesis. *J. Biol. Chem.* **276**, 37569–37576
  54. Bradbury, E. M., Inglis, R. J., Matthews, H. R., Langan, T. A. (1974) Molecular basis of control of mitotic cell division in eukaryotes. *Nature* **249**, 553–556
  55. Sung, M. T., Harford, J., Bundman, M., and Vidalakas, G. (1977) Metabolism of histones in avian erythroid cells. *Biochemistry* **16**, 279–285
  56. Wagner, T. E., Hartford, J. B., Serra, M., Vandegrift, V., and Sung, M. T. (1977) Phosphorylation and dephosphorylation of histone V (H5): controlled condensation of avian erythrocyte chromatin. Appendix: Phosphorylation and dephosphorylation of histone H5. II. Circular dichroic studies. *Biochemistry* **16**, 286–290
  57. Poccia, D., Lieber, T., and Childs, G. (1989) Histone gene expression during sea urchin spermatogenesis: an *in situ* hybridization study. *Mol. Reprod. Dev.* **1**, 219–229
  58. Hill, C. S., Rimmer, J. M., Green, B. N., Finch, J. T., and Thomas, J. O. (1991) Histone-DNA interactions and their modulation by phosphorylation of -Ser-Pro-X-Lys/Arg- motifs. *EMBO J.* **10**, 1939–1948