



High mobility group proteins: the multifaceted regulators of chromatin dynamics

Rwitie Mallik¹ · Anindya Kundu¹ · Shubho Chaudhuri¹

Received: 15 October 2018 / Accepted: 28 November 2018 / Published online: 5 December 2018
© Archana Sharma Foundation of Calcutta 2018

Abstract

DNA super-coiling and architectural proteins are the key players that maintain the chromatin in its compact state. Genomic DNA needs to be packaged such that it takes minimum space and can simultaneously be accessed for various DNA dependent processes. Architectural proteins are instrumental in organizing the dynamic higher order chromatin structure by effectuating a concerted effort among themselves and other nuclear proteins across spatio-temporal scales. The regulation of these proteins and their interaction with DNA modify the cellular phenotype by the modulation of gene expression. This review focuses on the structure–function relationship of three broad families of High mobility groups (HMGs) of protein, namely HMGA, HMGN and HMG-Box which are major chromatin architectural components of the eukaryotes. These nuclear elements not only act as architectural proteins but also play a multifaceted role in chromatin dynamics by facilitating interaction with nucleosomes, nucleosome-remodeling machines, transcription factors and histones.

Keywords Chromatin remodeling · Architectural protein · Linker histone H1 · HMG-box · HMGA · HMGN

Introduction

In eukaryotes, nuclear DNA are packaged tightly to be accommodated in the small nucleus which is accessible to the DNA dependent processes and at the same time space effective [61]. This versatile packaging of DNA makes the architecture of the genome extremely fascinating and one of the most studied aspect of genome biology in the last few decades.

Primarily, this packaging starts with wrapping of DNA strand around the histone octamer involving 1.7 turns encompassing 147 bps. The arrays of nucleosomes that constitute “beads on a string arrangement” structure are further packaged into the 30 nm fibre by coiling around itself. The stability of the higher order chromatin structure needs

the involvement of histone H1 that binds the linker DNA between two nucleosomes [92]. Although the exact organization of H1 is not known, but it is well established that H1 globular domain contacts the DNA near the nucleosome dyad axis and the adjacent linker DNA, thus stabilizes the wrapped DNA-histone octamer structure [27, 100, 119]. The highly basic C-terminal domain of H1 protein actually interacts with the negatively charged linker DNA and facilitates the chromatin compaction. But binding to DNA is not solely determined by the distribution of positive charged residue at the C-terminal of histone H1, rather specific residues play an essential role in maintaining the specificity of the interaction [9, 63, 104]. Importance of H1 is further accentuated by the fact that loss of H1 results in change in gene expression by regulating the access of transcriptional regulators.

However, disruption of the local and higher order nucleosome structure is the primary requisite for proper access of DNA sequences by different nuclear machinery. This disruption can be of two types: (1) transient unwrapping of DNA at certain position followed by re-wrapping at the same entry/exit point and (2) migration of the nucleosome complex along the DNA involving simultaneous unwrapping and reestablishment of new DNA protein contact with a new sequence. Thus, DNA packaging begins with supercoiling and then aided by macromolecular crowding which in turn is

This article is dedicated to the memory of Profs. A.K. Sharma and Archana Sharma.

Rwitie Mallik and Anindya Kundu have contributed equally to this work.

✉ Shubho Chaudhuri
shubho@jcbose.ac.in

¹ Division of Plant Biology, Bose Institute, P1/12 CIT Scheme VII M, Kolkata 700054, India

stabilized by DNA binding proteins such as histones. These DNA binding proteins that help to stabilize the DNA structure within the cell form the ‘Architectural proteins’ that are grouped as wrappers, benders and bridgers [56]. One such modifier is the High Mobility Group (HMG) class of proteins which are well known for their roles as architectural DNA binder in the nucleus and mitochondria, as signaling regulators in the cytoplasm and as inflammatory cytokines in the extracellular space [5].

HMG protein superfamilies are one of the major groups of architectural protein present in the eukaryotic cells. The HMG terminology is coined due to their unusual solubility, along with smaller size and migration properties in comparison to chromatin proteins during gel electrophoresis [85]. Intriguingly, recent evidences of live cell imaging also revaluated their rapid migration properties within the nucleus where they play the role of an essential partner among the dynamic network of architectural proteins that structurally modulate the chromatin to affect the downstream DNA-dependent activities [5]. Being the second most abundant protein next to histones, in the last decade a lot of effort has been put forward to decipher its mode of action during eukaryotic development. This review is a concerted effort to summaries the recent advancement in deciphering the role of 3 groups of HMGs namely HMGA, HMGN and HMG-Box (HMGB) in animals and plants.

HMGA protein superfamily: an overview

The HMG proteins belong to three families (HMGA, HMGB and HMGN), while members of each family are structurally divergent they share significant functional similarity. They do not possess any intrinsic transcriptional activity but due to its ability to modify the chromatin structure it allows binding of transcription factors (TFs) at the promoter and enhancer sequences [85]. As a result, they are often regarded as architectural TFs and are classified into three families with systematic reference to the DNA binding domains they contain:

- *HMGA* proteins contain AT-hooks, nine amino acid segments that are unstructured in solution but bind AT-rich DNA stretches in the minor groove.
- *HMGB* proteins contain HMG-Boxes, 80 amino acid domains that bind into the minor groove of DNA with limited or no sequence specificity.
- *HMGN* proteins bind inside nucleosomes, between the DNA spires and the histone octamer.

HMGA superfamily

In mammals, HMGAs are coded by two genes namely HMGA1 and HMGA2. HMGA1 gives rise to splice variants HMGA1a, HMGA1b and HMGA1c in certain rare cases [18]. All of these HMGAs have a canonical DNA-binding domain that recognizes the palindromic amino acid motif called the ‘AT-hook’. Except for HMGA1c, all HMGA proteins contain three such short basic AT-hook motifs and a C-terminal acidic tail (Fig. 1). The amino acid sequence of the AT-hook motif is K/RXRGRP (X = glycine or proline) with positively charged residues on both sides and it specifically recognizes the minor groove of AT-rich DNA stretches and binds to nucleosomes in a cooperative manner [18, 50, 113]. A distinguishing feature of HMGA proteins is their disordered random coils form as free molecules and the ability to take any defined secondary structure in the DNA bound form. This uniqueness in intrinsic flexibility and the ability of disordered-to-ordered structural transition following substrate binding makes it an essential player in a wide variety of biological processes. Along with its ability to recognize the structure of the narrow minor groove of A/T-rich DNA, HMGA proteins can also effectively recognize and bind to DNAs other than B-form. List of such binding substrates include: synthetic four-way and three-way junctions [29, 30], bent and supercoiled DNAs [73], base-unpaired regions of A/T-rich DNA [53] and distorted or flexible regions of DNA on isolated nucleosome core particles [83, 87]. In case of nucleosomes, the HMGA can bind in an ATP-independent process where it can induce localized changes in the rotational setting of DNA on the surface of reconstituted core particles [87]. Domain swap experiments using hybrid recombinant proteins revealed that the AT-hooks regions are responsible for nucleosome binding. HMGA1a and HMGA1b are highly homologous to each other, with HMGA1a having 11 more amino acids in mammals [89]. Irrespective of their high level of sequence homology, they are functionally quite distinct. While overexpression of HMGA1b in the human breast epithelial cell line MCF7 resulted in rapid proliferation into a metastatic and highly malignant phenotype, overexpression of HMGA1a did not result in such abnormality [84]. Such functional discrepancy between these HMGAs is probably due to post transcriptional modifications or it could be due to variation in the spacing of AT-hook domain [16]. Unlike in animals, plant HMGA members contain a typical GH1 domain along with the canonical AT-hook motif [44]. Structurally these proteins have a highly conserved central globular domain (GH1), 2 less conserved unstructured tail fragments: a short (~ 20 aa) N-terminal domain (NTD) and a considerably longer (~ 100 aa) and highly positively

High Mobility Group A

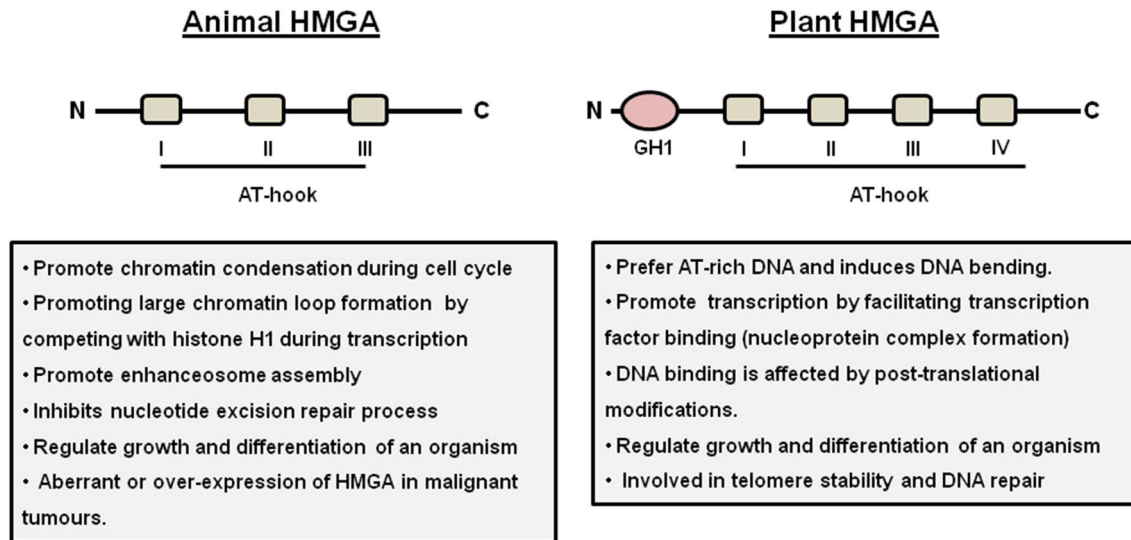


Fig. 1 Domain organization and salient features of High Mobility Group protein A (HMGA) of animal and plants. The grey box denotes AT-hook motif and purple circle denotes domain similar to

globular domain of histone H1 (GH1). GH1 domain is only found in plant HMGA family

charged C-terminal domain (CTD). The GH1 domain comprising of ~80 aa, belongs to the ‘winged helix’ family of DNA-binding proteins. As this fusion of GH1 domain to AT-hook motif is primarily found in angiosperms, evolutionary this protein structure might have evolved much later to the canonical HMGA proteins. In *Arabidopsis thaliana* there are 3 such proteins (GH1-HMGA1-3) that possess 4 to 6 AT-hook motifs. And interestingly, apart from plants fusion of GH1 and AT-hook is reported from primitive fish like *Trichoplax adhaerens* (the only extant representative of the phylum Placozoa), some yeast, nematode and insect species [44]. The proteins encoded by fish and *T. adhaerens* genomes are very large (up to 2900 amino acids) where GH1 and AT-hook motifs cooccur with RING and PHD domains.

Transcriptional regulon of HMGA

Inside the mammalian cell, HMGA1 proteins were found to co-localize with histone H1 at the scaffold attachment regions (SARs) [97]. HMGA proteins compete with histone H1 for binding to the linker DNA to destabilize the higher order chromatin structure and to make it accessible for transcription factor binding. As demonstrated by different biochemical techniques, this competition between HMGA1 and H1 resulted in loosening of the chromatin structure [13, 42, 118]. Post-translational modifications of H1 also play an important role in H1-HMGA interaction and exchange

[38]. Although, the exact mechanism of H1 replacement by HMGA is not known but similar post-translational modification dependent replacement is possible. Moreover, several motifs for kinases are also present inside the globular domain which interact with the linker DNA and phosphorylation of these sites affects its binding affinity [17], which might have a role during replacement of histone H1 by HMGA proteins. Such interactions of HMGA with different chromatin remodelers [59], possibly play an essential role in opening up the chromatin for the recruitment of TFs.

Transcription of genes are regulated by the core promoter elements in concert with regulatory elements like enhancers and silencers often present several base pairs upstream or downstream to the promoter sequence. During transcription, under the influence of defined signals, specific proteins bind to the enhancer sequence and form a complex called enhanceosome [66, 76, 117]. HMGA proteins by their intrinsic ability to bend DNA causes looping that bring the enhanceosome and the core promoter in close proximity resulting in the enhanced gene transcription [7, 15]. One of the best known examples is the regulation of interferon β (IFN- β) promoter where HMGA binding introduce bending of the DNA causing association of GCN5 (General Control Non-repressible 5), a histone acetyl transferase. GCN5 binds to this enhanceosomes and acetylate histones without altering their position helping binding of SWI/SNF factor which in turn shifts the second nucleosome 37 bps downstream to expose the TATA box [1, 117]. Thus, HMGA proteins

through DNA bending and by formation of enhanceosome complex cause chromatin rearrangements so that it can provide a conducive environment at the transcription start site (TSS) for the assembly of transcription initiation complex. Similar complex formation is also reported in plants, where rice HMGA protein PFI stimulates binding of transcription activator GT-2 to *PHYA* (phytochrome A) gene promoter to regulate its activity in a light dependent manner [102]. HMGA like protein NAT1 and LAT1 isolated from leaf and nodule nuclei, interact with different AT motifs in soybean nodulin promoter [36]. Recent study by Charbonnel et al. [14], has identified that a telomeric interacting protein called GH1-HMGA1 is involved in telomere stability and DNA repair. The *A. thaliana gh1-hmga1* mutant showed developmental and growth defects like, increased telomere instability, increased mitotic anaphase bridges, and higher sensitivity to DNA damaging agents like mitomycin-C and γ -irradiation.

HMG-Box protein super family

HMG-box domain containing proteins are the largest sub-family of high mobility group proteins that not only plays an essential role as architectural protein in DNA-dependent processes but also acts as an extracellular cytokine and an important component of autophagy pathway. HMG-Box is primarily comprised of 75aa containing domain which is ubiquitously found in all eukaryotes and was first reported to be present in high mobility group proteins, HMGB. Structural studies have shown HMG-box forms a L-shaped structure comprising of 3 α -helices which is conserved among all HMGBs irrespective of their amino acid sequence homology [102, 105]. The long arm consists of helix III and the N-terminal extended strand, whereas the short arm of the L-shape is composed of helices I and II at $\sim 80^\circ$ angle between them. The HMG-box binds the minor groove of the DNA where the hydrophobic residues of the L-shape intercalate between the DNA bases leading to widening of minor groove and unwinding of the DNA [41, 100]. HMG-box shows both sequence specific DNA binding (mammalian TFs such as SEX DETERMINING REGION OF Y [SRY] and LYMPHOID ENHANCER-BINDING FACTOR1 [LEF-1]) as well as non-specific DNA binding (chromosomal HMGB proteins and Structure-Specific Recognition Protein1 [SSRP1]), but their affinity for certain DNA structures like four-way junctions and DNA minicircles are noteworthy [10, 102, 105, 110].

DNA binding properties of HMG-box protein

In vertebrate, HMG1 and HMG2 are the two major HMG-box proteins that contain two HMG-box domains (A and

B) in tandem and a long acidic tail in the C-terminal end (Fig. 2). Both the A and B domain of HMG protein have relatively similar folded structure with subtle change in the orientation of helix I and II and the loop region between helix I and II [57, 111]. However, both the domains can bind with the DNA through minor groove. While domain A prefer to bind distorted DNA structure, domain B binding to DNA introduces almost 90° bend into the structure. The primary hydrophobic residue (Phe) present in the hydrophobic wedge of the DNA binding surface of B-type domain intercalates the DNA minor groove which produces a kink in the bound DNA backbone resulting in the widening of the minor group [103]. The second kink was introduced two bases away from the primary kink by the intercalation of second hydrophobic residue (Ile) in the minor groove. The basic extension present in many HMG-box proteins stabilizes this structure by binding to the compressed major groove present opposite to the widened minor groove. Domain A on the other hand does not have the primary intercalating residue; as a result can produce less bending angle compared to B domain after binding to the DNA [105]. Single HMG box domain containing proteins also shows sequence specific DNA binding except for few like HMGD of *Drosophila*, NHP6A of *Saccharomyces cerevisiae*. In these cases the DNA binding mechanism was very similar to that of HMG-box B.

In contrast to HMGB1 and HMGB2, mitochondrial transcription factor A (TFAM), which contains two tandem HMG-box domains, show both sequence specific and non-specific DNA binding properties [57]. In contrast to HMG-box domain of HMGB1, the primary intercalating residue of TFAM is non-polar whereas the second interacting residue is a polar residue that forms hydrogen bond with the DNA. The box A domain of TFAM has Leu as primary residue that intercalates DNA and the polar residue Thr as second intercalating residue which forms the hydrogen bond. The box on the other hand has an inverted motif with Asn as primary intercalation site that forms hydrogen bond and Leu as second intercalating residue. Together these two HMG-boxes of TFAM forms an “inverted tail to tail” type of configuration that create an 180° bend in the DNA [71, 96]. The overall bended structure is supported by the formation of additional α -helix in the linker region (region between the two HMG-box domains) that binds the minor groove of the DNA to neutralize the negative charge of the DNA backbone. Additional information from the crystal structure of TFAM-promoter DNA sequence reveals that the sequence specific binding of TFAM is governed by the complex network of interaction of the two HMG-boxes and the linker region with the DNA, that create additional sequence specific contact points to stabilizes the highly bent conformation of the DNA.

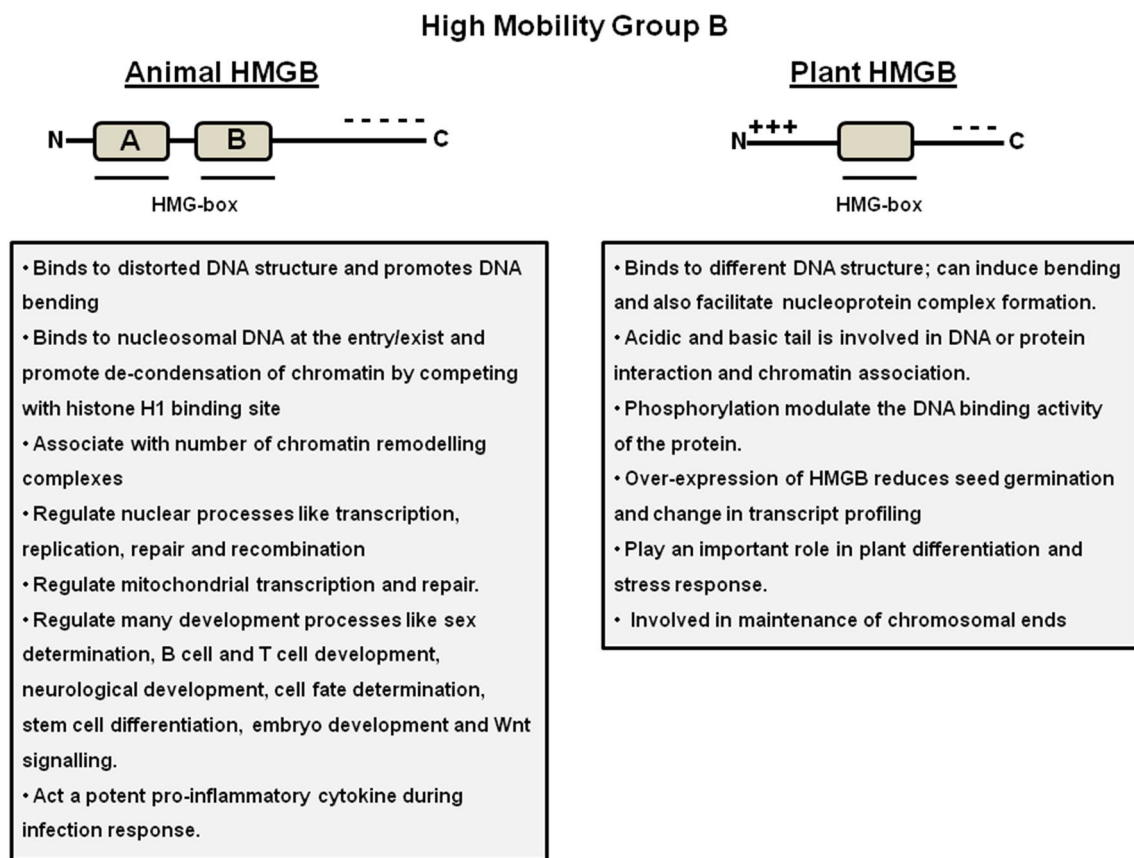


Fig. 2 Domain organization and salient features of High Mobility Group protein B (HMGB) of animal and plants The grey box denotes HMG-box motif. The basic and the acidic region is denoted by (+) and (–)

HMGB-box protein and transcription regulation

The architectural activity of HMG-box protein makes it a good candidate to interact with various sequence specific TFs to form a ternary complex with DNA [57]. In most of the cases, the role of HMG-box proteins is to pre-bend the DNA structure to favor the binding of TFs. HMGB1 has been shown to facilitate the binding of TF p53 by providing a favorable DNA bent structure as substrate [60, 94]. The ability of HMGB1 protein to bind distorted DNA structures in vivo provides an important feature that also facilitates transcription process. The class I steroid receptors has been shown to bind the DNA with less affinity and bend it moderately. The C-terminal extension of class I receptor recruit HMGB1 which binds this bended structure with high affinity and stabilizes the binding of class I receptor [62, 108]. Unlike HMGA, very few enhanceosomes formation was observed with HMGB: in BHLF-1 gene of Epstein-Barr virus, one enhanceosome is formed at the promoter and another on the enhancer. In this case, HMGB protein promotes the binding of a b-Zip protein called ZEBRA and Sp I to the DNA to form the enhanceosome which later recruit TFIIID and TFIIA [2, 20, 64]. In spite of having higher

abundance and efficient architectural function, the reason for having low HMGB enhanceosome inside the cell is not well understood. One possible reason may be because of its dynamic nature, to bind with different DNA sequences and to interact with various nuclear factors it rarely stays within the complex.

HMGB-box proteins were shown to facilitate nucleosome remodeling presumably by interfering the DNA-protein interaction. HMGB1 in collaboration with ATP-dependent chromatin remodelers has been shown to promote nucleosome sliding. In this case, HMGB1 binds to the nucleosomal DNA at the entry/exit point and create a bend structure. This bend structure is stabilized by the basic region of the HMGB protein that neutralizes the negative charge of the sugar-phosphate backbone of the opposite strand. The acidic tail of HMGB further stabilizes the positive charges of histones thereby decrease the affinity of the DNA towards histone core [54, 78, 105]. The distorted DNA structure facilitates the binding of chromatin remodeler ACF, which slides the DNA bulge around the surface of nucleosome core. In the whole process it was believe that formation of initial bend by HMGB1 is the rate limiting step. Further involvement of the HMG proteins with the chromatin modulation comes

from the fact that it is found to be associated with a number of remodeling complexes such as SWI/SNF complexes of *Drosophila* (BRM, *brahma* complex) [77], mammalian SWI/SNF complex BAF complex and histone chaperone FACT (facilitates chromatin transcription) complex [109].

Other than acting as transcription activator, HMG-box protein has been shown to regulate transcription by acting as transcription repressor. The best example in this case is the regulation of Wnt signaling by HMG protein. The Wnt signaling pathway is an important signal network that regulates cell fate determination, cell migration, cell polarity, stem cell pluripotency and organogenesis during embryonic development [74]. Studies have shown that the misregulation of Wnt signal can cause various diseases including cancer. The T cell factor/lymphoid enhancer-binding factor (TCF/LEF) family of proteins is important HMG domain containing proteins that play a role in the Wnt signaling cascade. TCF/LEF proteins bind to the promoters of Wnt target genes and keep them shut down in the absence of signaling cues. In the absence of Wnt signaling, β -catenin is captured by the destruction complex [Axin and adenomatous polyposis coli (APC), the Ser/Thr kinases GSK-3 and CK1, protein phosphatase 2A (PP2A), and the E3-ubiquitin ligase β -TrCP] and is phosphorylated for subsequent proteosomal degradation [101]. As the Wnt signaling pathway is turned on, the β -catenin is released from the destruction complex and cytosolic β -catenin builds up. The cytosolic β -catenin then shuttles to the nucleus and interacts with the TCF/LEF proteins where it removes the co-repressors and brings in the co-activators to turn on the transcription of Wnt target genes [12]. The transcriptional repression role of HMG-box protein in Wnt signaling has also been demonstrated in lower organisms like *Xenopus*, *Drosophila* and *Caenorhabditis elegans* [43]. The *Xenopus* HMG-box proteins SOX17a/b and the SOX3 proteins physically interact with β -catenin and repress the Wnt signaling. Human HBP1 is a HMG-box protein that acts as a cell cycle repressor [98]. Studies have shown that HBP1 binds to the promoter of Cyclin D1 and inhibits the expression of the gene, thereby functioning as transcriptional repressors of Wnt signaling.

HMG proteins promote de-condensation of chromatin by antagonizing the role of H1

About 147 bp of DNA wraps around a nucleosome 1.7 times. The first and last point of contact between the nucleosome and the DNA wrapping are known as the entry and exit site of the nucleosome which mark the most accessible regions of the nucleosome due to the transient association with DNA at these sites [28]. The H1 and HMG proteins have similar DNA binding properties and bind at entry or exit site of the nucleosome and at linker DNA dyad [3, 72]. While H1 proteins were known for chromatin compaction and stabilisation

of 30 nm fibre, HMG proteins were shown to compete with H1 to facilitate chromatin opening and remodeling [92]. Thus the H1 histones and the HMGs play an antagonistic role in the chromatin dynamics. This reciprocal relationship is best elucidated by the studies on embryogenesis of *Xenopus* oocytes and *Drosophila* [70, 72]. While H1 levels are barely detectable at the early stages, with progress in development the expression of H1 increases and it can also replace HMGs in some places. For programmed cellular development selective silencing of gene locus is an important aspect. Experimental evidences in *Drosophila* have shown that antagonistic binding of HMG and H1 to genomic locations can bring about transcriptional changes of various gene loci [69]. Studies have also shown that HMG and H1 proteins compete for specific DNA binding sites and HMGs can weaken the H1 binding considerably.

Plant HMG-box proteins

While human genome encodes for 47 HMG-box proteins having molecular weight ranging from 15 to 193 kD, higher plant genomes encode for only 10 to 15 different HMG-box proteins of approximately 13 to 72 kD [102]. While in humans, HMG-box domain was found in many TFs which represent the largest subgroup, in plants no such TFs have been reported which clearly indicates that HMG-box TFs are more diversified in mammalian system in comparison to land plants [110]. Phylogenetic analysis of all the HMG-box containing proteins present in land plants and primitive species like *Selaginella moellendorffii* (Pteridophyte), the moss *Physcomitrella patens*, and algae like *Chlamydomonas reinhardtii* and *Volvox carteri* indicate them to be classified into four distinct families: chromosomal HMGB proteins, AT-rich interaction domain (ARID)-HMG proteins, 3xHMG-box proteins and SSRP1 [4].

Chromosomal HMGB proteins

Plants express a higher number of HMGB variants in comparison to other eukaryotes and they are structurally different from mammalian HMGB proteins [102]. While animal HMGB proteins have two HMG-boxes with an intermediate basic linker and C-terminal flanking acidic region, plant HMGB proteins have one HMG-box flanked with N-terminal basic and C-terminal acidic region (Fig. 2). A search in database indicates that all land plants to code for HMGBs and *A. thaliana* genome codes for 8 such proteins [4]. The DNA binding studies have indicated that both the N-terminal basic domain (which increases DNA binding) and the acidic C-terminal domain (which reduces DNA binding) regulate the DNA-protein interaction [91]. Like animal counterpart, plant HMGB proteins also binds to different DNA topological structures and can produce bend in the DNA backbone.

Although there is no clear evidence of plant HMGB mediated enhanceosome formation, but like mammalian HMGB, plant HMGB has been shown to interact with TFs like DOF (DNA binding with one finger). Interestingly, the prior interaction between DOF2 and HMGB facilitate the DNA binding of DOF2 protein whereas CK2 mediated phosphorylation of HMGB1 abolishes the interaction and negatively regulates the DNA binding [99, 115].

While mammalian HMGB1 is reported to be located outside nucleus acting as cytokines [65, 116], plant HMGB proteins like AtHMGB1/5 are mostly nuclear localized [24, 85]. Moreover, efficient nuclear localization of AtHMGB1/5 requires the basic N-terminal region whereas C-terminal acidic region interfere with nuclear targeting. Immunolocalization of AtHMGB1 in the meristematic root tip cells shows spotted distribution pattern associated with interphase chromatin but not with condensed mitotic chromosomes. Fluorescence recovery after photobleaching experiments revealed highly dynamic nature of AtHMGB1/5 protein in the nucleus where they bind chromatin transiently before moving to the next binding site. Apart from AtHMGB1/5, other HMGB protein such as AtHMGB2/3 and AtHMGB4 were found to be shuttling between the nucleus and cytosol [80]. The nucleo-cytoplasmic distribution of these groups of protein depends upon the amino acid sequence of the N-terminal basic and C-terminal acidic region of the protein, where C-terminal acidic tail is largely responsible for the shuttling of the protein between nucleus and cytoplasm. Currently it is unclear why HMGB shuttles between nucleus and cytoplasm in plants and there is no evidence for the occurrence of HMG protein in plant mitochondria or chloroplast.

Exposure of plants to environmental signals require a switch in the gene expression program which is needed for immediate stress response and later on for the adaptability of plant to environmental conditions. The induction of these arrays of stress response genes requires changes in the chromatin structure to facilitate the machinery to initiate transcription. It has been found that the expression of *Arabidopsis* HMGB genes is differentially regulated by abiotic stress treatment [40, 47], indicating their essential role in modulating chromatin structure to facilitate gene expression [37, 47]. Transcriptomic profiling of *hmg1* revealed a large number of cell cycle-related factors downregulated in comparison to *col-0* plants, which is in concert with reduced root length in *hmg1* mutants [52]. Moreover, several salt stress-responsive genes were also downregulated in *hmg1* mutants as compared to control plants, indicating a relationship between HMGB driven expression of genes under salinity stress.

HMGB protein is also found to play an important role in plant differentiation. Cells having reduced expression of cotton HMGB3 gene which is preferentially expressed in embryonic tissue show altered potential of differentiation

and dedifferentiation during somatic embryogenesis [32]. These cells having reduced expression of HMGB3 show differential expression of genes involved in pathway similar to β -catenin signaling. Recently a new role of HMGB protein has been identified in the maintenance of chromosomal ends where *A. thaliana* mutant *hmg1* shows shortened telomeres and plants overexpressing HMGB1 have elongated telomeres [82]. Although there is no change in the activity of telomerase in these plants, it is possible that HMGB1 influence the telomere chromatin structure in order to maintain the chromosomal ends.

AT-rich interaction domain ARID-HMG proteins

ARID-HMG is a unique group of plant proteins belonging to the HMG-box superfamily having two DNA binding domains, one N-terminal ARID and a C-terminal HMG-box, which together codes for a 34–56 kD protein (except for *Physcomitrella* where it is 82 kD) [4, 90] (Fig. 3). Uniquely this family of proteins are absent in mammalian system and among the angiosperms are more diversified in dicots (sixteen) in comparison to monocots (five). Comparison at the amino acid level shows ARID-HMG proteins can be phylogenetically classified into 4 subgroups. Of these four subgroups, AtARID-HMG1/2 belonging to two different subgroups was found to be more widely expressed than the other ARID-HMGs and found to be nuclear localized in BY-2 protoplasts [26]. Although independently both ARID and HMG domain are reported for DNA binding capability, recent in silico docking simulation for AtHMGB11 indicated ARID domain is specifically responsible for DNA binding [95]. ARID-HMG proteins prefer to bind AT-rich sequence in comparison to GC-rich sequence and can also binds to different DNA structures like supercoiled and mini-circles [10, 25, 26]. Interestingly, these proteins can also induce negative supercoiling in the relaxed plasmid and can bend the DNA [95]. In *A. thaliana* HMGB15 interacts with TFs AGL66 and AGL104 in vitro and *hmg15* mutants showed delayed pollen tube germination indicating its role in seed development [114] Comparative transcriptomic analysis between *hmg15* and WT (*Col-0*) revealed that genes involved in cell wall synthesis, osmoregulation, solute transporter, lipid synthesis, carbohydrate synthesis and stress response were transcriptionally affected in the mutant. Extensive analysis showed *Lotus japonicus* ARID protein SIP1 bind AT-rich elements of the NIN promoter and has been suggested to play an essential role during Rhizobium-legume plant symbiosis [120]. However, whether there is any role of ARID-HMG protein in plant-bacteria symbiotic pathway is not yet known.

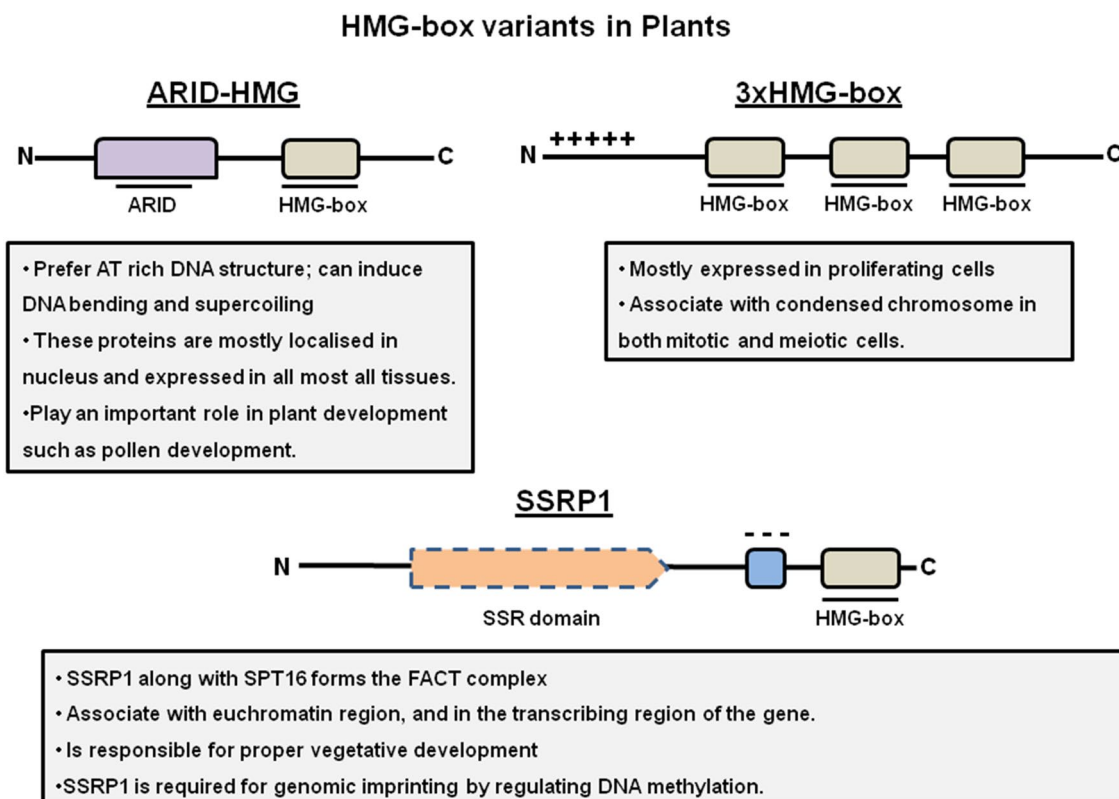


Fig. 3 Domain organization and salient features Plant HMG-box variants The grey box denotes HMG-box motif

3xHMG-box proteins

3xHMG-box codes for protein of 43–60 kD with unique N-terminal basic domain followed by three HMG-box domain in tandem (Fig. 3). Like ARID-HMG, 3xHMG class of proteins is exclusively found in plants and unlike ARID-HMG is equally represented in monocots and dicots [4]. In Antosch et al. [4] except *A. thaliana* and *Populus trichocarpa* where two 3xHMGs are found, all other plants species have only one gene coding for 3xHMG protein. An exclusive phylogenetic analysis of only HMG-box domains from different HMG proteins including 3xHMGs indicated that irrespective of their functional similarity they have diverged uniquely. In contrast to other HMG-box proteins, 3xHMG-box proteins are widely expressed in the plant, specifically in the proliferative mitotic cells. Immunofluorescence studies for At3xHMG-box1 and At3xHMG-box2 indicated their association with condensed chromosomes during different stages of M-phases [79]. Unlike other HMG group proteins that specifically bind interphase chromatin, 3xHMG always associate with mitotic chromosomes [49, 79]. Even in meiotic cells, 3xHMG-box proteins were found to interact with condensed chromosomes in pollen mother cells [79]. This close association of 3xHMGs during mitosis and meiosis lead to the

proposition that they are involved in chromosomal condensation during segregation.

SSRP

SSRP1 is the fourth type of chromatin modifiers from the HMG-box superfamily which functions with another protein SPT16 by forming dimeric facilitates chromatin transcription (FACT) complex [8, 75]. FACT complex, first identified in yeast and mammalian system, behaves like a histone chaperone by assisting RNA Polymerase II during transcription elongation by initiating nucleosome disassembly. Interestingly, FACT complex are also responsible for re-assembling of chromatin post transcription and thus maintain a homeostasis by preventing cryptic transcript initiation [21, 88, 112].

SSRP1 is a highly conserved protein coded by a single gene in most flowering plants and as well as in *Selaginella*, *Physcometrella* and algae. Plant SSRP1 is of 61–78 kD and shows overall structural similarity to animal SSRP1 with homology in N-terminal, middle acidic and C-terminal HMG-box domain but unlike animal SSRP1 lacks a C-terminal tail of 80 amino acid [93] (Fig. 3). Evidences suggest that in maize, N-terminal of HMG-box is responsible for nuclear localization of SSRP1. Like any other HMGs,

HMG-box is responsible for recognition of DNA, but unlike other HMGs, SSRP1 recognition of DNA is sequence independent. Moreover, SSRP1 recognizes nucleosome particles and the super coiled and minicircle DNA by its structure where the specificity of DNA binding is regulated through CK2 dependent phosphorylation of SSRP1 [45, 51, 93]. *A. thaliana* mutants of SSRP1 resulted in various defects in vegetative and generative development with marked increase in number of leaves, inflorescence showing early bolting and lack of seed production [55].

Role of SSRP in genome imprinting

Another interesting aspect of SSRP1 is its role in parental-specific trans-generational memory through genome imprinting. SSRP1 is responsible for DNA demethylation which causes repression of parentally imprinted genes in the female central cell before fertilization [35]. Although lack of demethylation in *ssrp1-3* mutant resulted in decrease in expression of maternally expressed genes, paternally inherited genes showed upregulation defying the usually known mechanism of methylation dependent gene silencing. Similar observations were made for maternal alleles of HDG3 and VIM5 that are upregulated in both PRC2 and DNA methylase mutants [31]. Most of the functional aspect of SSRP1 is studied in association with SPT16 in FACT (facilitates chromatin transactions) complex, where it is involved in the DME (DEMETER)-dependent regulation of genomic imprinting in *A. thaliana* endosperm [35]. Recent evidences show that DME in association to FACT complex is responsible for genome wide DNA-demethylation of GC-rich heterochromatin domains with high nucleosome occupancy enriched with H3K9me2 and H3K27me1 [22].

HMGN superfamily

HMGN is a group of high mobility proteins that are exclusively found in mammalian system which can specifically bind nucleosome to induce chromatin modifications and epigenetic changes [46]. HMGN protein family is comprised of five proteins with a conserved nucleosomal binding domain (NBD), unique bipartite nuclear localization signal (NLS) at the N terminal, a nucleosomal binding domain and an acidic C-terminal chromatin regulatory domain (CHUD) [11] (Fig. 4). The nucleosome binding domain of HMGN contains a conserved sequence RRSARLSA that promote the binding of the protein to the nucleosome [107]. Sequence analysis of HMGN1-4 shows that they are small proteins with ~90 aa whereas HMGN5 has unique structure with unusual long C-terminal region (~200 aa residue in human) [23]. In vitro and in vivo studies show that all five HMGNs can bind to chromatin with similar affinity and the binding

Nucleosome Binding High Mobility Group (HMGN)

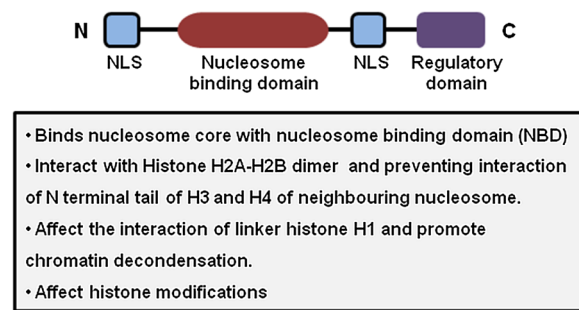


Fig. 4 Domain organization and salient features of animal nucleosome binding High Mobility Group protein (HMGN)

is not dependent on the DNA sequence. None of the plants have yet been found to code for any HMGN homologue.

Recent results from methyl-TROSY NMR spectroscopy indicate that one molecule of HMGN protein binds to each side of the nucleosome core through Nucleosome binding domain (NBD) [39]. The conserved sequence of NBD interacts with the acidic patch formed by group of non-histone H2A-H2B dimer. The C-terminal region of NBD interacts with DNA near the entry/exit region of nucleosome core facilitating the C-terminal region of HMGN to interact with the linker DNA by interfering with the binding of proteins that can specifically bind inside nucleosome between histone H1. Surprisingly, HMGNs do not displace the H1 from nucleosomes and instead bind simultaneously with H1 [67]. They however directly alter both, the nucleosome-dependent condensation of the H1 CTD and core histone tail interactions. HMGN1 and HMGN2 reduce the propensity of nucleosome arrays to undergo self-association into higher order chromatin structures, in an H1-dependent manner. In case of HMGN5, which has unusual longer negative charged C-terminal region, interact with the positive charged H1 more efficiently than other HMGN [58]. HMGN not only interact with histone H1 but also affect the interaction of N termini of histone H3 and H4 with the neighboring nucleosome. Thus by interfering with the binding of histone H1 and affecting the inter-nucleosomal interaction, HMGN reduces the compaction of the chromatin to increase the accessibility of nucleosomal DNA to different regulatory factors of different DNA-dependent processes.

HMGN regulates epigenetic modifications of chromatin landscape

Apart from the architectural nature of HMGN in altering the chromatin structure, evidence has shown its role in modulating histone modifications. Since HMGN interact with nucleosome and interact with different tail residues in

the nucleosome, it is likely that histone modifications can be affected by this interaction. HMGN1 has been shown to enhance H3K14 acetylation, acetylation and methylation of H3K9 residue, phosphorylation of H3S10 and H2AS1 [46]. HMGN induced H3K14ac at the promoter region of Hsp70 gene is an important epigenetic modification for its transcription, as *hmg1* plants showed lowered transcript density and H3K14ac marks HMGN1 binding to nucleosome has been shown to reduce the H3S10 phosphorylation by interfering with the ability of the kinase to phosphorylate H3. H3S10 phosphorylation is an important epigenetic mark that alters chromatin structure during many cellular processes such as transcription activation to chromatin condensation during mitosis [46]. Recent study has shown that HMGN in embryonic stem cells (ESC) of mouse plays an important role in their differentiation along the neuronal pathway. During stem cell differentiation the loss of HMGN1 affects the expression of two transcription factor OLIG1 and OLIG2 required for oligodendrocyte lineage specification [19]. Loss of HMGN1 lead to increase binding of histone H1 and increase of H3K27me3 repressive mark at Olig1 and Olig2 genes leading to transcriptional repression. Genome-wide studies have shown that HMGN1 occupancy in the genome overlaps with DNase I hypersensitive sites which includes promoter, enhancer and transcription factor binding site. It will be interesting to investigate whether HMGN1 binding and its influence to modulate histone modifications promote and maintain the DHS sites in the chromatin.

Role of High Mobility Group protein proteins in DNA repair process

Efficient repair of damaged DNA is a major challenge in eukaryotic cells because of the constrain imposed due to complex chromatin structure. The major pathway of repair from UV damage and bulky adducts from DNA is by the nucleotide excision repair (NER) system. Studies have revealed that recovery rate from UV damage is faster in naked DNA in comparison to nucleosome bound DNA [68]. Also disruption of histones and DNA contact or change in global compaction of chromatin can lead to faster UV repair [48]. The inability of the repair machinery to repair the damage site can lead to diseases such as xeroderma pigmentosum, trichothiodystrophy.

HMGN proteins have been shown to enhance NER in the context of chromatin. HMGN proteins affect the stability of the higher order chromatin structure by targeting histone H1 and the H3 N terminal tail leading to a decompaction of structure [106]. Studies have shown that HMGN1 mutant mice and mouse embryonic fibroblasts are more sensitive to UV irradiation in comparison to normal. The sensitivity can be returned to normalcy upon the expression of wild type HMGN1 in the mutant [6]. The study proposed that HMGN1

reduces the compaction of chromatin fibers in irradiated cells, and facilitates the accessibility of NER machinery to repair the DNA lesion caused by UV damage.

HMGA and HMGB have been shown to behave differently during damage repair response in contrast to HMGN. HMGB1 preferentially binds to damaged DNA isolated from cisplatin treated cells that induces 1,2-intrastrand d(GpG) and d(ApG) cross-links [34, 81]. This cisplatin induced cross linking induces a bend in the DNA backbone. Studies have shown that HMGB1 binds to the cisplatin induced bended structure and inhibit the repair process by preventing the accessibility of NER proteins [33]. Similar to HMGB protein, HMGA proteins have been shown to inhibit the UV-induced CPDs (Cyclobutane pyrimidine dimer) probably because of tight binding of the protein to distorted DNA that shield the damage from repair processes [86]. Since plant genome does not code for HMGN, it will be interesting to investigate whether plant HMGA or HMG-box proteins plays an important role in DNA damage accessibility, repair process and subsequent restoration of chromatin structure.

Outlook

The ubiquitous presence of high mobility group proteins inside the cell suggests that they are involved in important biological function. By virtue of having versatile DNA binding ability, these groups of proteins have widely been reported to regulate various DNA dependent processes inside the nucleus that involves transcription regulation, chromatin remodeling, DNA replication, repair and recombination. Recent studies in animal and plants have demonstrated additional role of these proteins inside the cell such as genomic imprinting, chromosome condensation and segregation. Despite numerous studies in animal as well as in plant, the exact biological function of architectural protein remains an important area of research in chromatin biology. Especially for plants, the presence of diverse HMG-box proteins is a unique feature. Plant genome does not code for any HMGN group of protein. The question remains, whether any of these HMG-box proteins can substitute the role of HMGN in plants. Another important aspect of plant architectural proteins is to understand the role of these HMG proteins in modulating the epigenetic language of cell. This is also an unexplored territory which needs further investigation. Recent advancement in genomics and proteomics along with genetic studies in animal and plant can provided us with novel insight into their mode of actions.

Acknowledgements The authors sincerely apologize to all the authors whose work has not been acknowledged and referred in the review article. This work was supported by CSIR, Government of India; Grant Numbers [38(1245)/10/EMR-II] and [38(1365)/13/EMR-II] and core

grant of Bose Institute, Kolkata, India. A.K. also acknowledges DBT/RA for fellowship.

References

- Agalioti T, Lomvardas S, Parekh B, Yie J, Maniatis T, Thanos D. Ordered recruitment of chromatin modifying and general transcription factors to the IFN- β promoter. *Cell*. 2000;103:667–78.
- Agresti A, Bianchi ME. HMGB proteins and gene expression. *Curr Opin Genet Dev*. 2003;13:170–8.
- An W, van Holde K, Zlatanova J. The non-histone chromatin protein HMG1 protects linker DNA on the side opposite to that protected by linker histones. *J Biol Chem*. 1998;273:26289–91.
- Antosch M, Mortensen SA, Grasser KD. Plant proteins containing high mobility group box DNA-binding domains modulate different nuclear processes. *Plant Physiol*. 2012;159:875–83.
- Bianchi ME, Agresti A. HMG proteins: dynamic players in gene regulation and differentiation. *Curr Opin Genet Dev*. 2005;15:496–506.
- Birger Y, West KL, Postnikov YV, Lim JH, Furusawa T, Wagner JP, Laufer CS, Kraemer KH, Bustin M. Chromosomal protein HMG1 enhances the rate of DNA repair in chromatin. *EMBO J*. 2003;22:1665–75.
- Bouallaga I, Massicard S, Yaniv M, Thierry F. An enhanceosome containing the Jun B/Fra-2 heterodimer and the HMG-I (Y) architectural protein controls HPV18 transcription. *EMBO Rep*. 2000;1:422–7.
- Brewster NK, Johnston GC, Singer RA. Characterization of the CP complex, an abundant dimer of Cdc68 and Pob3 proteins that regulates yeast transcriptional activation and chromatin repression. *J Biol Chem*. 1998;273:21972–9.
- Brown DT, Izard T, Misteli T. Mapping the interaction surface of linker histone H1 with the nucleosome of native chromatin in vivo. *Nat Struct Mol Biol*. 2006;13:250.
- Bustin M. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol*. 1999;19:5237–46.
- Bustin M. Chromatin unfolding and activation by HMGN(*) chromosomal proteins. *Trends Biochem Sci*. 2001;26:431–7.
- Cadigan KM, Waterman ML. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Perspect Biol*. 2012;4:a007906.
- Catez F, Yang H, Tracey KJ, Reeves R, Misteli T, Bustin M. Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. *Mol Cell Biol*. 2004;24:4321–8.
- Charbonnel C, Rymarenko O, Da Ines O, Benyahya F, White CI, Butter F, Amiard S. The linker histone GH1-HMGA1 is involved in telomere stability and DNA damage repair. *Plant Physiol*. 2018;177:01789.
- Chen B, Young J, Leng F. DNA bending by the mammalian high-mobility group protein AT hook 2. *Biochemistry*. 2010;49:1590–5.
- Cleynen I, Van de Ven WJ. The HMGA proteins: a myriad of functions. *Int J Oncol*. 2008;32:289–305.
- Contreras A, Hale TK, Stenoien DL, Rosen JM, Mancini MA, Herrera RE. The dynamic mobility of histone H1 is regulated by cyclin/CDK phosphorylation. *Mol Cell Biol*. 2003;23:8626–36.
- Cui T, Leng F. Specific recognition of AT-rich DNA sequences by the mammalian high mobility group protein AT-hook 2: a SELEX study. *Biochemistry*. 2007;46:13059–66.
- Deng T, Postnikov Y, Zhang S, Garrett L, Becker L, Rácz I, Hölter SM, Wurst W, Fuchs H, Gailus-Durner VJ. Interplay between H1 and HMG1 epigenetically regulates OLIG1&2 expression and oligodendrocyte differentiation. *Nucleic Acids Res*. 2017;45:3031–45.
- Ellwood K, Huang W, Johnson R, Carey MJM. Multiple layers of cooperativity regulate enhanceosome-responsive RNA polymerase II transcription complex assembly. *Mol Cell Biol*. 1999;19:2613–23.
- Formosa T. The role of FACT in making and breaking nucleosomes. *Biochim Biophys Acta (BBA) Gene Regul Mech*. 2012;1819:247–55.
- Frost JM, Kim MY, Park GT, Hsieh P-H, Nakamura M, Lin SJ, Yoo H, Choi J, Ikeda Y, Kinoshita T. FACT complex is required for DNA demethylation at heterochromatin during reproduction in Arabidopsis. *Proc Natl Acad Sci USA*. 2018;115:E4720–9.
- Furusawa T, Cherukuri S. Developmental function of HMG1 proteins. *Biochim Biophys Acta (BBA) Gene Regul Mech*. 2010;1799:69–73.
- Grasser KD, Launholt D, Grasser M. High mobility group proteins of the plant HMGB family: dynamic chromatin modulators. *Biochim Biophys Acta (BBA) Gene StructExpr*. 2007;1769:346–57.
- Grosschedl R, Giese K, Pagel J. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet*. 1994;10:94–100.
- Hansen FT, Madsen CK, Nordland AM, Grasser M, Merkle T, Grasser KD. A novel family of plant DNA-binding proteins containing both HMG-box and AT-rich interaction domains. *Biochemistry*. 2008;47:13207–14.
- Hayes JJ, Pruss D, Wolffe AP. Contacts of the globular domain of histone H5 and core histones with DNA in a “chromatosome”. *Proc Natl Acad Sci*. 1994;91:7817–21.
- Hieb AR, Gansen A, Böhm V, Langowski J. The conformational state of the nucleosome entry–exit site modulates TATA box-specific TBP binding. *Nucleic Acids Res*. 2014;42:7561–76.
- Hill DA, Pedulla ML, Reeves R. Directional binding of HMG-I (Y) on four-way junction DNA and the molecular basis for competitive binding with HMG-1 and histone H1. *Nucleic Acids Res*. 1999;27:2135–44.
- Hill DA, Reeves R. Competition between HMG-I (Y), HMG-1 and histone H1 on four-way junction DNA. *Nucleic Acids Res*. 1997;25:3523–31.
- Hsieh T-F, Shin J, Uzawa R, Silva P, Cohen S, Bauer MJ, Hashimoto M, Kirkbride RC, Harada JJ, Zilberman D. Regulation of imprinted gene expression in Arabidopsis endosperm. *Proc Natl Acad Sci*. 2011;108:1755–62.
- Hu L, Yang X, Yuan D, Zeng F, Zhang X. GhHmgB3 deficiency deregulates proliferation and differentiation of cells during somatic embryogenesis in cotton. *Plant Biotechnol J*. 2011;9:1038–48.
- Huang J-C, Zamble DB, Reardon JT, Lippard SJ, Sancar A. HMG-domain proteins specifically inhibit the repair of the major DNA adduct of the anticancer drug cisplatin by human excision nuclease. *Proc Natl Acad Sci*. 1994;91:10394–8.
- Hughes EN, Engelsberg BN, Billings PC. Purification of nuclear proteins that bind to cisplatin-damaged DNA. Identity with high mobility group proteins 1 and 2. *J Biol Chem*. 1992;267:13520–7.
- Ikeda Y, Kinoshita Y, Susaki D, Ikeda Y, Iwano M, Takayama S, Higashiyama T, Kakutani T, Kinoshita T. HMG domain containing SSRP1 is required for DNA demethylation and genomic imprinting in Arabidopsis. *Dev Cell*. 2011;21:589–96.
- Jacobsen K, Laursen NB, Jensen EO, Marcker A, Poulsen C, Marcker KA. HMG I-like proteins from leaf and nodule nuclei interact with different AT motifs in soybean nodulin promoters. *Plant Cell*. 1990;2:85–94.
- Jang JY, Kwak KJ, Kang H. Expression of a high mobility group protein isolated from *Cucumis sativus* affects the germination

- of *Arabidopsis thaliana* under abiotic stress conditions. *J Integr Plant Biol.* 2008;50:593–600.
38. Ju B-G, Lunyak VV, Perissi V, Garcia-Bassets I, Rose DW, Glass CK, Rosenfeld MG. A topoisomerase II β -mediated dsDNA break required for regulated transcription. *Science.* 2006;312:1798–802.
 39. Kato H, van Ingen H, Zhou BR, Feng H, Bustin M, Kay LE, Bai Y. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc Natl Acad Sci USA.* 2011;108:12283–8.
 40. Kim JM, To TK, Nishioka T, Seki MJP. Chromatin regulation functions in plant abiotic stress responses. *Plant Cell Environ.* 2010;33:604–11.
 41. King CY, Weiss MA. The SRY high-mobility-group box recognizes DNA by partial intercalation in the minor groove: a topological mechanism of sequence specificity. *Proc Natl Acad Sci USA.* 1993;90:11990–4.
 42. Kishi Y, Fujii Y, Hirabayashi Y, Gotoh Y. HMGA regulates the global chromatin state and neurogenic potential in neocortical precursor cells. *Nat Neurosci.* 2012;15:1127.
 43. Kormish JD, Sinner D, Zorn AM. Interactions between SOX factors and Wnt/ β -catenin signaling in development and disease. *Dev Dyn.* 2010;239:56–68.
 44. Kotliński M, Knizewski Ł, Muszewska A, Rutowicz K, Lirski M, Schmidt A, Baroux C, Ginalski K, Jerzmanowski A. Phylogeny-based systematization of *Arabidopsis* proteins with histone H1 globular domain. *Plant Physiol.* 2017;174:00214.
 45. Krohn NM, Stemmer C, Fojan P, Grimm R, Grasser KD (2003) Protein kinase CK2 phosphorylates the HMG domain protein SSRP1 inducing the recognition of UV-damaged DNA. *J Biol Chem.* 2003;278:12710–5.
 46. Kugler JE, Deng T, Bustin M. The HMGN family of chromatin-binding proteins: dynamic modulators of epigenetic processes. *Biochim Biophys Acta.* 2012;1819:652–6.
 47. Kwak KJ, Kim JY, Kim YO, Kang H. Characterization of transgenic *Arabidopsis* plants overexpressing high mobility group B proteins under high salinity, drought or cold stress. *Plant Cell Physiol.* 2007;48:221–31.
 48. Lans H, Marteijn JA, Vermeulen WJE. ATP-dependent chromatin remodeling in the DNA-damage response. *Epigenetics Chromatin.* 2012;5:4.
 49. Launholt D, Merkle T, Houben A, Schulz A, Grasser KD. *Arabidopsis* chromatin-associated HMGA and HMGB use different nuclear targeting signals and display highly dynamic localization within the nucleus. *Plant Cell.* 2006;18:2904–18.
 50. Li O, Vasudevan D, Davey CA, Dröge P. High-level expression of DNA architectural factor HMGA2 and its association with nucleosomes in human embryonic stem cells. *Genesis.* 2006;44:523–9.
 51. Lichota J, Grasser KD. Differential chromatin association and nucleosome binding of the maize HMGA, HMGB, and SSRP1 proteins. *Biochemistry.* 2001;40:7860–7.
 52. Lildballe DL, Pedersen DS, Kalamajka R, Emmersen J, Houben A, Grasser KD. The expression level of the chromatin-associated HMGB1 protein influences growth, stress tolerance, and transcriptome in *Arabidopsis*. *J Mol Biol.* 2008;384:9–21.
 53. Liu W-M, Guerra-Vladusic FK, Kurakata S, Lupu R, Kohwi-Shigematsu T. HMG-I (Y) recognizes base-unpairing regions of matrix attachment sequences and its increased expression is directly linked to metastatic breast cancer phenotype. *Can Res.* 1999;59:5695–703.
 54. Lnenicek-Allen M, Read CM, Crane-Robinson C. The DNA bend angle and binding affinity of an HMG box increased by the presence of short terminal arms. *Nucleic Acids Res.* 1996;24:1047–51.
 55. Lolas IB, Himanen K, Grønlund JT, Lynggaard C, Houben A, Melzer M, Van Lijsebettens M, Grasser KD. The transcript elongation factor FACT affects *Arabidopsis* vegetative and reproductive development and genetically interacts with HUB1/2. *Plant J.* 2010;61:686–97.
 56. Luijsterburg MS, White MF, van Driel R, Dame RT. The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol.* 2008;43:393–418.
 57. Malarkey CS, Churchill ME. The high mobility group box: the ultimate utility player of a cell. *Trends Biochem Sci.* 2012;37:553–62.
 58. Malicet C, Rochman M, Postnikov Y, Bustin M. Distinct properties of human HMGN5 reveal a rapidly evolving but functionally conserved nucleosome binding protein. *Mol Cell Biol.* 2011;31:2742–55.
 59. Malini E, Maurizio E, Bembich S, Sgarra R, Edomi P, Manfioletti G. HMGA Interactome: new insights from phage display technology. *Biochemistry.* 2011;50:3462–8.
 60. McKinney K, Prives CJM. Efficient specific DNA binding by p53 requires both its central and C-terminal domains as revealed by studies with high-mobility group 1 protein. *Mol Cell Biol.* 2002;22:6797–808.
 61. Mekhail K, Moazed D. The nuclear envelope in genome organization, expression and stability. *Nat Rev Mol Cell Biol.* 2010;11:317.
 62. Melvin VS, Roemer SC, Churchill ME, Edwards DP. 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.* 2002;277:25115–24.
 63. Misteli T, Gunjan A, Hock R, Bustin M, Brown DT. Dynamic binding of histone H1 to chromatin in living cells. *Nature.* 2000;408:877.
 64. Mitsouras K, Wong B, Arayata C, Johnson RC, Carey M. The DNA architectural protein HMGB1 displays two distinct modes of action that promote enhanceosome assembly. *Mol Cell Biol.* 2002;22:4390–401.
 65. Müller S, Ronfani L, Bianchi M. Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function. *J Intern Med.* 2004;255:332–43.
 66. Munshi N, Yie J, Merika M, Senger K, Lomvardas S, Agaloti T, Thanos D. The IFN- β enhancer: a paradigm for understanding activation and repression of inducible gene expression. In: Cold Spring Harbor symposia on quantitative biology, vol. 64. Cold Spring Harbor Laboratory Press; 1999, pp. 149–160.
 67. Murphy KJ, Cutter AR, Fang H, Postnikov YV, Bustin M, Hayes JJ. HMGN1 and 2 remodel core and linker histone tail domains within chromatin. *Nucl Acids Res.* 2017;45:9917–30.
 68. Nag R, Smerdon MJ. Altering the chromatin landscape for nucleotide excision repair. *Res Rev Mutat Res.* 2009;682:13–20.
 69. Nalabothula N, McVicker G, Maiorano J, Martin R, Pritchard JK, Fondufe-Mittendorf YN. The chromatin architectural proteins HMGD1 and H1 bind reciprocally and have opposite effects on chromatin structure and gene regulation. *BMC Genom.* 2014;15:92.
 70. Ner SS, Blank T, Perez-Paralle ML, Grigliatti TA, Becker PB, Travers AA. HMG-D and histone H1 interplay during chromatin assembly and early embryogenesis. *J Biol Chem.* 2001;276:37569–76.
 71. Ngo HB, Kaiser JT, Chan DC. The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nat Struct Mol Biol.* 2011;18:1290.
 72. Nightingale K, Dimitrov S, Reeves R, Wolffe AP. Evidence for a shared structural role for HMGI and linker histones B4 and H1 in organizing chromatin. *EMBO J.* 1996;15:548–61.
 73. Nissen MS, Reeves R. Changes in superhelicity are introduced into closed circular DNA by binding of high mobility group protein I/Y. *J Biol Chem.* 1995;270:4355–60.

74. Nusse R, Clevers HJC. Wnt/ β -catenin signaling, disease, and emerging therapeutic modalities. *Cell*. 2017;169:985–99.
75. Orphanides G, Wu W-H, Lane WS, Hampsey M, Reinberg D. The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature*. 1999;400:284.
76. Panne D. The enhanceosome. *Curr Opin Struct Biol*. 2008;18:236–42.
77. Papoulas O, Daubresse G, Armstrong JA, Jin J, Scott MP, Tamkun JW. The HMG-domain protein BAP111 is important for the function of the BRM chromatin-remodeling complex in vivo. *Proc Natl Acad Sci USA*. 2001;98:5728–33.
78. Payet D, Travers A. The acidic tail of the high mobility group protein HMG-D modulates the structural selectivity of DNA binding 1. *J Mol Biol*. 1997;266:66–75.
79. Pedersen DS, Coppens F, Ma L, Antosch M, Markt B, Merkle T, Beemster GT, Houben A, Grasser KD. The plant-specific family of DNA-binding proteins containing three HMG-box domains interacts with mitotic and meiotic chromosomes. *New Phytol*. 2011;192:577–89.
80. Pedersen DS, Merkle T, Markt B, Lildballe DL, Antosch M, Bergmann T, Tönsing K, Anselmetti D, Grasser KD. Nucleocytoplasmic distribution of the Arabidopsis chromatin-associated HMGB2/3 and HMGB4 proteins. *Plant Physiol*. 2010;154:110.163055.
81. Pil PM, Lippard SJ. Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science*. 1992;256:234–7.
82. Prochazkova Schrupfova P, Fojtova M, Mokros P, Grasser KD, Fajkus JJ. Role of HMGB proteins in chromatin dynamics and telomere maintenance in Arabidopsis thaliana. *Curr Protein Pept Sci*. 2011;12:105–11.
83. Pruss D, Reeves R, Bushman FD, Wolffe AP. The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J Biol Chem*. 1994;269:25031–41.
84. Reeves R. Molecular biology of HMGA proteins: hubs of nuclear function. *Gene*. 2001;277:63–81.
85. Reeves R. Nuclear functions of the HMG proteins. *Biochim Biophys Acta*. 2010;1799:3–14.
86. Reeves R, Adair JE. Role of high mobility group (HMG) chromatin proteins in DNA repair. *DNA Repair*. 2005;4:926–38.
87. Reeves R, Nissen M. Interaction of high mobility group-I (Y) nonhistone proteins with nucleosome core particles. *J Biol Chem*. 1993;268:21137–46.
88. Reinberg D, Sims RJ. de FACTo nucleosome dynamics. *J Biol Chem*. 2006;281:23297–301.
89. Resar LM. The high mobility group A1 gene: transforming inflammatory signals into cancer? *Cancer Res*. 2010;70:436–9.
90. Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science*. 2000;290:2105–10.
91. Ritt C, Grimm R, Fernandez S, Alonso JC, Grasser KD. Basic and acidic regions flanking the HMG domain of maize HMGA modulate the interactions with DNA and the self-association of the protein. *Biochemistry*. 1998;37:2673–81.
92. Robinson PJ, Fairall L, Huynh VA, Rhodes DJ. EM measurements define the dimensions of the “30-nm” chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci*. 2006;103:6506–11.
93. Röttgers K, Krohn NM, Lichota J, Stemmer C, Merkle T, Grasser KD. DNA-interactions and nuclear localisation of the chromosomal HMG domain protein SSRP1 from maize. *Plant J*. 2000;23:395–405.
94. Rowell JP, Simpson KL, Stott K, Watson M, Thomas JO. HMGB1-facilitated p53 DNA binding occurs via HMG-Box/p53 transactivation domain interaction, regulated by the acidic tail. *Structure*. 2012;20:2014–24.
95. Roy A, Dutta A, Roy D, Ganguly P, Ghosh R, Kar RK, Bhunia A, Mukhobadhyay J, Chaudhuri S. Deciphering the role of the AT-rich interaction domain and the HMG-box domain of ARID-HMG proteins of Arabidopsis thaliana. *Plant Mol Biol*. 2016;92:371–88.
96. Rubio-Cosials A, Sydow JF, Jiménez-Menéndez N, Fernández-Millán P, Montoya J, Jacobs HT, Coll M, Bernadó P, Solà M. Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat Struct Mol Biol*. 2011;18:1281.
97. Saitoh Y, Laemmli UK. Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell*. 1994;76:609–22.
98. Sampson EM, Haque ZK, Ku MC, Tevosian SG, Albanese C, Pestell RG, Paulson KE, Yee ASJTEj (2001) Negative regulation of the Wnt- β -catenin pathway by the transcriptional repressor HBP1. 20: 4500-4511.
99. Schultz TF, Spiker S, Quatrano RS. Histone H1 enhances the DNA binding activity of the transcription factor EmBP-1. *J Biol Chem*. 1996;271:25742–5.
100. Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu R-M, Zhu P, Li G. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science*. 2014;344:376–80.
101. Stamos JL, Weis WI. The β -catenin destruction complex. *Cold Spring Harbor Perspect Biol*. 2013;5:a007898.
102. Stros M, Launholt D, Grasser KD. The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell Mol Life Sci*. 2007;64:2590–606.
103. Stros M, Muselikova E. A role of basic residues and the putative intercalating phenylalanine of the HMG-1 box B in DNA supercoiling and binding to four-way DNA junctions. *J Biol Chem*. 2000;275:35699–707.
104. Th’ng JP, Sung R, Ye M, Hendzel MJ. H1 family histones in the nucleus control of binding and localization by the C-terminal domain. *J Biol Chem*. 2005;280:27809–14.
105. Thomas JO, Travers AA. HMGI and 2, and related ‘architectural’ DNA-binding proteins. *Trends Biochem Sci*. 2001;26:167–74.
106. Trieschmann L, Postnikov YV, Rickers A, Bustin M. Modular structure of chromosomal proteins HMG-14 and HMG-17: definition of a transcriptional enhancement domain distinct from the nucleosomal binding domain. *Mol Cell Biol*. 1995;15:6663–9.
107. Ueda T, Catez F, Gerlitz G, Bustin M. Delineation of the protein module that anchors HMGN proteins to nucleosomes in the chromatin of living cells. *Mol Cell Biol*. 2008;28:2872–83.
108. Verrijdt G, Haelens A, Schoenmakers E, Rombauts W, Claessens F. Comparative analysis of the influence of the high-mobility group box 1 protein on DNA binding and transcriptional activation by the androgen, glucocorticoid, progesterone and mineralocorticoid receptors. *Biochem J*. 2002;361:97–103.
109. Wang W, Chi T, Xue Y, Zhou S, Kuo A, Crabtree GR. Architectural DNA binding by a high-mobility-group/kinesin-like subunit in mammalian SWI/SNF-related complexes. *Proc Natl Acad Sci*. 1998;95:492–8.
110. Wegner M. All purpose Sox: the many roles of Sox proteins in gene expression. *Int J Biochem Cell Biol*. 2010;42:381–90.
111. Weir HM, Kraulis PJ, Hill CS, Raine AR, Laue ED, Thomas JO. Structure of the HMG box motif in the B-domain of HMGI. *EMBO J*. 1993;12:1311–9.
112. Winkler DD, Luger K. The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization. *J Biol Chem*. 2011;R110:180778.

113. Winter N, Nimzyk R, Bösche C, Meyer A, Bullerdiek J. Chromatin immunoprecipitation to analyze DNA binding sites of HMGA2. *PLoS One*. 2011;6:e18837.
114. Xia C, Wang YJ, Liang Y, Niu QK, Tan XY, Chu LC, Chen LQ, Zhang XQ, Ye D. The ARID-HMG DNA-binding protein AtHMGB15 is required for pollen tube growth in *Arabidopsis thaliana*. *Plant J*. 2014;79:741–56.
115. Yanagisawa S. Dof DNA-binding domains of plant transcription factors contribute to multiple protein-protein interactions. *Eur J Biochem*. 1997;250:403–10.
116. Yang H, Tracey KJ. Targeting HMGB1 in inflammation. *Biochim Biophys Acta (BBA) Gene Regul Mech*. 2010;1799:149–56.
117. Yie J, Merika M, Munshi N, Chen G, Thanos D. The role of HMG I (Y) in the assembly and function of the IFN- β enhanceosome. *EMBO J*. 1999;18:3074–89.
118. Zhao K, Käs E, Gonzalez E, Laemmli UK. SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J*. 1993;12:3237–47.
119. Zhou B-R, Feng H, Kato H, Dai L, Yang Y, Zhou Y, Bai Y. Structural insights into the histone H1-nucleosome complex. *Proc Natl Acad Sci*. 2013;110:19390–5.
120. Zhu H, Chen T, Zhu M, Fang Q, Kang H, Hong Z, Zhang Z. A novel ARID DNA-binding protein interacts with SymRK and is expressed during early nodule development in *Lotus japonicus*. *Plant Physiol*. 2008;148:337–47.