Cell. Mol. Life Sci. 64 (2007) 2590 – 2606
1420-682X/07/192590-17
DOI 10.1007/s00018-007-7162-3

Cellular and Molecular Life Sciences

Review

The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins

M. Štros[a, *], D. Launholt[b] and K. D. Grasser[b, *]

[a] Laboratory of Analysis of Chromosomal Proteins, Academy of Sciences of the Czech Republic, Institute of Biophysics, Královopolská 135, 612 65 Brno (Czech Republic), e-mail: stros@ibp.cz
[b] Department of Life Sciences, Aalborg University, Sohngaardsholmsvej 49, 9000 Aalborg (Denmark), e-mail: kdg@bio.aau.dk

Received 30 March 2007; received after revision 29 May 2007; accepted 18 June 2007
Online First 29 June 2007

Abstract. The HMG-box domain of ~75 amino acid residues was originally identified as the domain that mediates the DNA-binding of chromatin-associated high-mobility group (HMG) proteins of the HMGB type. In the last few years, HMG-box domains have been found in various DNA-binding proteins including transcription factors and subunits of chromatin-remodeling complexes. HMG-box domains mediate either non-sequence-specific (*e.g.*, HMGB-type proteins) or sequence-specific (*e.g.*, transcription factors) DNA binding. Both types of HMG-box domains bind non-B-type DNA structures (bent, kinked and unwound) with high affinity. In addition, HMG-box domains are involved in a variety of protein-protein interactions. Here, we have examined the human and plant genomes for genes encoding HMG-box domains. Compared to plants, human cells contain a larger variety of HMG-box proteins. Whereas in humans transcription factors are the most divergent group of HMG-box proteins, in plants the chromosomal HMGB-type proteins are most variable.

Keywords. High-mobility group (HMG) protein, DNA binding, chromatin, architectural factor, transcription factor, genomic stability, human and plant genome.

Introduction

The eukaryotic DNA is very long and must be highly condensed to fit into the cell nucleus. The first level of compactness involves coiling of the DNA around histone octamers [consisting of the central $(H3/H4)_2$ tetramer and two peripheral H2A/H2B dimers] to form nucleosomes [1]. The nucleosomes are further arranged with the assistance of linker histones to form higher order 30-nm chromatin fibers comprising the chromosomes. Modulation of chromatin folding affects access of regulatory factors to their cognate DNA binding sites by remodeling of the chromatin structure. This can be achieved by loosening the chromatin structure, and in some cases by disruption of the nucleosome structure, by DNA bending and unwinding, as well as by affecting the strength of DNA-histone interactions by post-synthetic modifications of histones.

Many of these structural changes are mediated by a large and diverse superfamily of high mobility group (HMG) proteins. Three structurally distinct classes of HMG proteins have been defined: the HMG-nucleosome binding family (HMGN), the HMG-AT-hook family (HMGA), and the HMG-box family (HMGB)

* Corresponding authors.

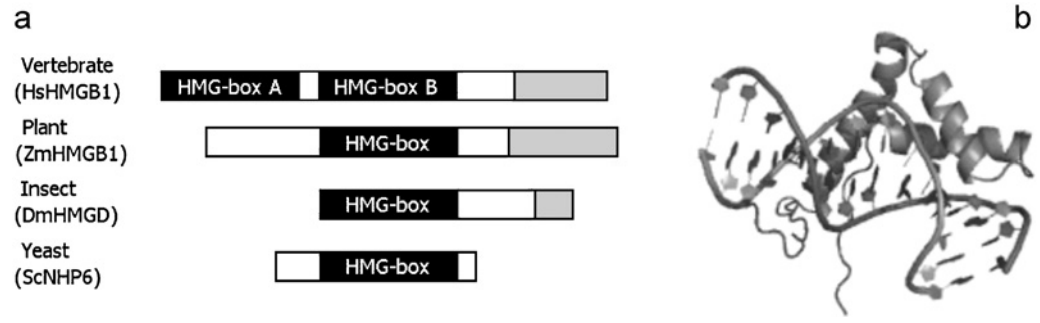

Figure 1. (*a*) Schematic representation of structures of HMGB-type proteins from various eukaryotes. HMG-boxes are indicated by black boxes, while the C-terminal acidic domains found in the plant, insect and vertebrate proteins, but not in the yeast NHP6A/B proteins, are indicated by dark gray boxes. HMG-boxes of the "single HMG-box proteins" have a higher degree of similarity to the second HMG-box (domain B) than to the first HMG-box (domain A) of mammalian HMGB1. It should be pointed out that additional putative HMGB-type proteins exist that have different overall structures such as the *Drosophila* DSP-1 and the yeast HMO1 proteins (both having two HMG-box domains) [18, 19]. (*b*) Solution NMR structure of the HMG-box domain of the human male sex-determining factor SRY complexed to DNA [13]. The three α-helices of the HMG-box domain and their binding within the minor groove of DNA is depicted. The SRY-DNA structure was retrieved from Swiss-Prot database (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1j46).

[2–4]. Mammalian HMG-box containing proteins are usually classified into two major groups. The first group consists of HMGB-type non-sequence-specific DNA binding proteins with two HMG-box domains and a long highly acidic C-tail (see Fig. 1a). Since the discovery of similar DNA-binding motifs within the nucleolar transcription factor UBF [5], numerous other HMG-box proteins have been identified [6, 7]. This second group is highly diverse and consists of proteins having mostly a single HMG-box and no acidic C-tails, but there are also proteins that have up to six copies of the HMG-box domain (Table 1).

The solution structures of individual HMG boxes (domains A and B) of HMGB1 have been determined by NMR spectroscopy [8–10]. The HMG-box domain of ~75 amino acids has a characteristic L-shaped fold consisting of three α-helices with an angle of ~80° between the arms [3]. The long arm includes the extended N-terminal strand and helix III, while the short arm is composed of helixes I and II. The overall structure of the HMG-box is far more conserved than the corresponding amino acid sequences of the different HMG-boxes [11]. The HMG-boxes of mammalian HMGB1 (domains A and B) differ slightly in the orientation of helices I and II with respect to the rest of the molecule, and also the length of the loop between helices I and II is longer in the domain A [10]. The structures of the HMG-boxes have been determined from a number of non-sequence-specific and sequence-specific HMG-box proteins, and although all are very similar to the structures of the two HMG-boxes from HMGB1, their structures resemble most closely domain B. Several structures of single HMG-boxes in complex with linear DNA (LEF-1 [12]; SRY box [13], see Fig. 1b; yeast NHP6A [14]; *Drosophila* HMG-D [15]), or with bent, cisplatin-modified, DNA (domain A of HMGB1 [16]) have been determined, and the structures of bound and free HMG-boxes are similar. In mammals, HMGB1 and other related HMGB-type proteins typically contain two HMG-boxes that recognize distorted DNA structures with very weak or no sequence specificity, while transcription factors of the HMG-box family contain only one HMG-box that binds both structure and sequence specifically to DNA. Binding of HMG-boxes to the DNA minor groove causes unwinding and widening of the minor groove accompanied by bending. The extent of DNA bending varies among HMG-boxes and depends on sequence differences on only very few positions {*e.g.*, 54° for SRY (see Fig. 1b) and 110° for LEF-1 due to partial intercalations of Ile and Met of helices I of the HMG-boxes, respectively [12, 13]}. In each case the HMG-box protein binds on the outside of the DNA bend to compress the major groove. Recently, the first structure of two tandem HMG-boxes (an engineered version of the AB di-domain of HMGB1, in which the domain A has been replaced by the HMG-box of the sequence-specific transcription factor SRY, to give SRY.B) bound to a short linear DNA was presented [17], indicating that the linker between the two HMG-boxes is positioned within the minor groove.

In this review, we focus on plant and mammalian (mostly human) non-sequence-specific and sequence-specific HMG-box proteins with the main emphasis on their possible nuclear functions. Although human and plant HMG-box proteins have a different organization outside the HMG-box(es), the plant and human HMG-box domains are remarkably conserved, suggesting that they are involved in similar biological functions in these evolutionarily distant organisms. The analysis of phylogenetic trees of plant and human HMG-box proteins revealed the different

Table 1. Human HMG-box proteins.

	Protein name	Swiss-Prot	Gene locus	Length (aa)	Mass (Da)	HMG-box	Function
	HMGB1	P09429	13q12	215	24894	2	Architectural protein, genome stability, signal transduction
	HMGB2	P26583	4q31	209	24034	2	Architectural protein, similar to HMGB1, male fertility
	HMGB3	O15347	Xq28	200	22980	2	Protein deficiency results in erythrocythemia
	HMG1L10	Q9UGV6	22q12.1	211	24218	2	Unknown
	Sp100-HMG	P23497	2q37.1	879	100417	2	Repressor/(or activator) of viral and cellular promoters
	HMG4L	Q3SYE8	20q11.22	130	14606	1	Unknown
	HMG2L1	Q9UGU5	22q13.1	601	65712	1	Transcriptional repressor of Wnt signaling
	SSRP1	Q08945	11q12	709	81075	1	A component of the transcription elongation FACT complex
	HMGB4	Q8WW32	1p35.1	186	22405	2	Unknown
	TFAM	Q00059	10q21	246	29097	2	Mitochondrial transcription factor
	UBF1	P17480	17q21.3	764	89406	6	RNA polymerase I transcription factor
	LEF1	Q9UJU2	4q23-q25	399	44201	1	Regulate gene expression during cell differentiation
	TCF1	P36402	5q31.1	383	41686	1	Effector of Wnt signaling directing cell fate determination
	TCF3	Q9HCS4	2p11.2	588	62631	1	Repressor of of Wnt signaling directing cell fate determination
	BBX	Q8WY36	3q13.1	941	105130	1	Transcription factor necessary for progression from G1 to S phase
	BRAF35	Q9P0W2	19p13.3	317	35813		Chromatin remodeling factor
	HMG20A	Q9NP66	15q24	347	40144		Chromatin remodeling factor
	PMS1	P54277	2q31.1	932	105830		Mismatch repair protein
	CAGF9	O15405	16q12.1	331	37161	1	T lymphocyte-specific factor in the thymus development
	GCX1	Q96NM4	20q13.12	488	51604	1	T lymphocyte-specific factor in the thymus development
	LCP1	O94842	14q11.2	621	66195	1	T lymphocyte-specific factor in the thymus development
	TOX	O94900	8q12.1	526	57513	1	T lymphocyte-specific factor in the thymus development
A	**SRY**	Q05066	Yp11.3	204	23884	1	Sex-determining factor, necessary for testes development
	SOX1	O00570	13q34	387	38856	1	Suppress neurogenesis
B1	**SOX2**	P48431	3q26.3-q27	317	34310	1	Suppress neurogenesis
	SOX3	P41225	Xq27.1	446	45210	1	Suppress neurogenesis
B2	**SOX14**	O95416	3q22-q23	240	26485	1	Cell-type specification along the dorsoventral axis
	SOX21	Q9Y651	13q31-q32	276	28580	1	Promotes neuronal differentiation
	SOX4	Q06945	6p22.3	474	47263	1	Transcriptional activator in neuronal maturation
C	**SOX11**	P35716	2p25	441	46679	1	Transcriptional activator in neuronal maturation
	SOX12	O15370	20p13	315	34122	1	Unknown
	SOX5	P35711	12p12.1	763	84026	1	Regulate oligodendrocyte development, opposite to SOX9/10
D	**SOX6**	P35712	11p15.3	828	91893	1	Regulate oligodendrocyte development, opposite to SOX9/10
	SOX13	Q9UN79	1q32	889	98782	1	Involved in chondrogenesis, neurogenesis, and limb development
	SOX8	P57073	16p13.3	446	47314	1	Neural crest development
E	**SOX9**	P48436	17q24.3q25.1	509	56137	1	Major downstream effector of SRY, involved in testis development
	SOX10	P56693	22q13.1	466	49911	1	Formation of neural crest cells, specification of derivative cell fates
	SOX7	Q9BT81	8p22	388	42197	1	Transcriptional repressor of Wnt signaling
F	**SOX17**	Q9H6I2	8q11.23	414	44117	1	Specification of cardiac mesoderm in embryonic stem cells
	SOX18	P35713	20q13.33	384	40891	1	Regulator of blood vessel formation
G	**SOX15**	O60248	17p13	233	25251	1	Involved in trophoblast giant cell differentiation
H	**SOX30**	O94993	5q33	753	81854	1	Transcriptional factor involved in development
	BAF 57	Q969G3	17q21.2	411	46649	1	Chromatin remodeling factor
	CIC	Q96RK0	19q13.2	1608	163820	1	Involved in granule cell development
	HBP1	O60381	7q22-q31	514	57645	1	Inhibits G and-S1-phase progression, suppressor of Wnt signaling
	PB1	Q86U86	3p21	1689	192948	1	A component of a chromatin remodeling complex PBAF
	WHSC1	O96028	4p16.3	1365	152258	1	Possibly involved in chromosomal translocations in myelomas

groups of HMG-box proteins that occur in mammals and plants.

Human proteins containing HMG-box domains

We have used the amino acid sequence of the human HMGB1 HMG-box di-domain (A and B domain, residues 1–185) as query to identify HMG-box proteins in the human protein database (http://www.ebi.ac.uk/swissprot/) using BLASTP. The search revealed a total of 47 amino acid sequences of 15–193 kDa, exhibiting sequence similarity to the HMG-box domains of the human HMGB1 protein. The retrieved sequences were aligned and used to construct an amino acid sequence similarity tree to

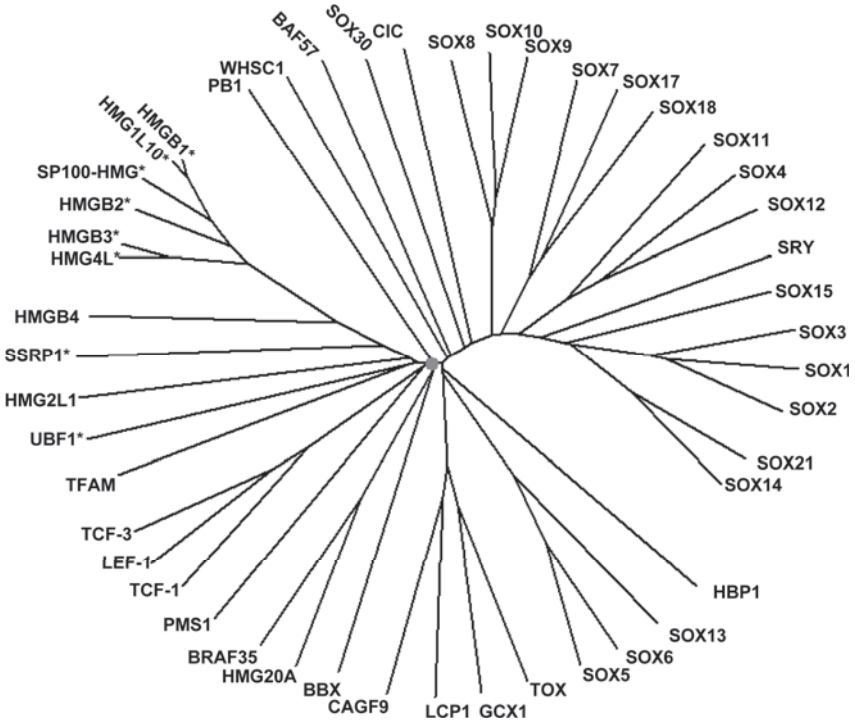

Figure 2. Unrooted phylogenetic tree of human HMG-box proteins. The amino acid sequence of human HMGB1 AB didomain (residues 1–185) was used as query to identify similar human HMG-box proteins in the Swiss-Prot database (http://www.ebi.ac.uk/swissprot/) using the protein-protein BLAST program (http://130.14.29.110/BLAST). The retrieved sequences (see Table 1 for the corresponding Swiss-Prot accession numbers) were aligned by multiple sequence alignment that was used to construct the amino acid sequence similarity tree using the ClustalX software (http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html). Sequences with asterisks indicate HMG-box proteins containing, in addition to the variable number of HMG-boxes, acidic C-terminal domains.

reveal a sequence relationship between the proteins. The human HMG-box proteins are also listed in Table 1, indicating their similarities, molecular masses, and possible functions. From the tree in Figure 2 follows that the human HMG-box proteins can be subdivided into seven groups. The first two groups contain variable numbers of HMG-boxes, whereas all other groups contain proteins with a single HMG-box. Group 1 consists of HMG-box proteins containing two HMG-boxes with the exception of a single HMG-box protein HMG4L of unknown function. This group comprises HMGB-type proteins with two HMG-boxes including HMGB1–3, transcriptional activator or repressor SP100-HMG (a splicing variant of SP100 containing an HMG-box domain) [20], and HMG1L10 of unknown function. Group 2 consists of five proteins related to group 1: single HMG-box containing transcriptional repressor HMG2L1 [21] and the structure-specific recognition protein SSRP1 [22, 23], mitochondrial transcription factor TFAM containing two HMG-boxes [24] and HMGB4 of unknown function, and upstream binding transcription factor UBF1 containing six HMG-boxes [25, 26]. The transcription factors TCF/LEF-1 (a group of proteins that regulate gene expression during cell differentiation) form group 3. Group 4 includes chromatin-remodeling factors HMG20A, BRAF35 (also termed HMG20B, [27]), mismatch repair protein PMS1 [28] and SOX-related protein BBX (bobby SOX homolog of a *Drosophila* protein). Group 5 represents the TOX family (T lymphocyte-specific factors) proteins implicated in the regulation of thymocyte selection [29]. The sex-determining factor SRY and the SOX (SRY-related) proteins are arranged into group 6. These proteins can be further divided into several distinct subgroups [30]: A (SRY), B1 (SOX1–3) and B2 (SOX14, 21), C (SOX4, 11, 12), D (SOX5, 6, 13), E (SOX8–10), F (SOX7, 17, 18), G (SOX15) and H (SOX30). Group 7 includes diverse HMG-box proteins: a SOX-related CIC protein [31], chromatin-remodeling factors BAF57 and PB1, a suppressor of Wnt signaling HBP1 [32] and protein WHSC1 (possibly involved in Wolf-Hirschhorn malformation syndrome, [33]).

In most cases, the domain B of HMGB1 is more closely related to HMG-boxes of other HMG-box proteins than the domain A of HMGB1 [3, 10], although we have noticed that 11 out of 47 amino acid sequences of human HMG-box proteins were only identified when either the domain A (or the AB didomain, Table 1) or the full-length HMGB1 [34] were used as query. These proteins included SRY, the F subgroup of SOX proteins, TFAM, CIC, BAF57, WHSC1, LEF-1, TCF-1 and TCF-2, indicating higher similarity of this group of HMG-box proteins to the domain A rather than to the domain B of HMGB1.

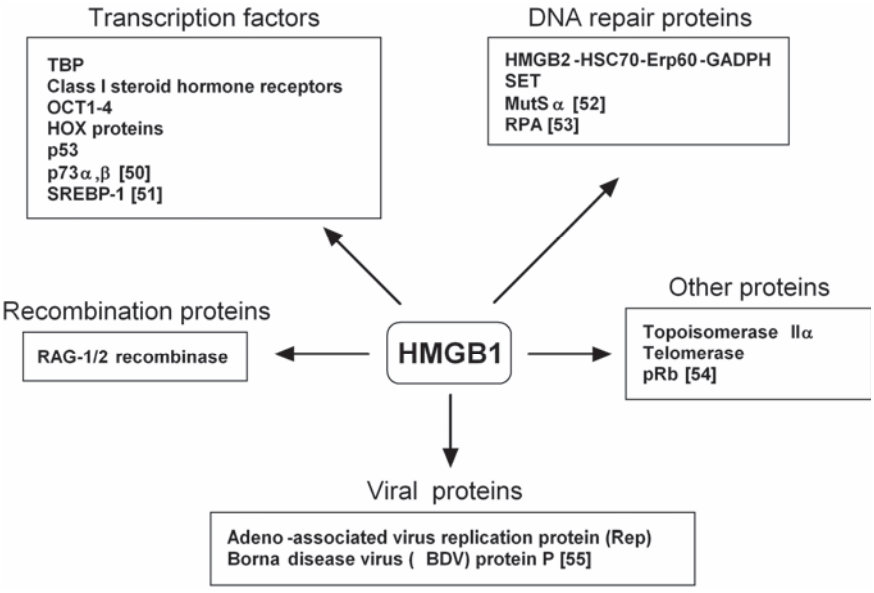

Figure 3. Protein binding partners of HMGB1. Most of the HMGB1 protein interactions have been determined by *in vitro* pull-down assays, and the depicted interacting proteins represent mainly DNA-binding proteins. With the exception of topoisomerase IIalpha (Štros et al., 2007, in press) and telomerase (Kunická et al., submitted), all HMGB1-interacting proteins or complexes that are indicated without references are to found in [2].

Human HMGB-type proteins

HMG proteins were originally classified by their anomalous electrophoretic mobility in Triton-urea gels (resulting from a high content of charged amino acid residues) and their extractability in 0.35 M NaCl and in diluted acids solutions [35]. There are three HMGB variants in human and mouse, HMGB1, HMGB2 and HMGB3. While HMGB1–3 proteins are expressed in early mouse embryos, HMGB2 and HMGB3 are down-regulated during embryonic development [36–40]. The abundant HMGB1 protein (approximately 1 molecule per 10–15 nucleosomes) is highly conserved among mammals, and it continues to be ubiquitously expressed in adults. Mammalian HMGB-type proteins (HMGB1–3) have a tripartite domain organization, consisting of two DNA-binding domains, the HMG-boxes A and B, and acidic C-terminal tails of variable length. While the HMGB1-boxes interact with DNA, the C-tail usually decreases the affinity of the protein for DNA [3, 41]. Apart from the canonical HMGB-type proteins HMGB1–3, the only HMG-box proteins containing acidic C domains are HMG4L, HMG1L10, SP100-HMG, UBF1 and SSRP-1, (Fig. 2, proteins indicated with asterisks). Thus, unlike the plant genomes where the majority of HMG-box proteins represent the HMGB-type (see below), human HMGB proteins form a minority among the HMG-box proteins (Fig. 2).

Most of the available data on possible functions of mammalian (human) HMGB-type proteins are related to HMGB1. HMGB-type proteins represent architectural factors ("DNA chaperones") in the chromatin facilitating formation of complex nucleoprotein structures [42]. The function of HMGB proteins is clearly dependent on their ability to interact with DNA (in particular with distorted or prebent DNA structures), and to induce specific changes in the target DNA structure by bending/looping and unwinding [41, 43–47]. The specific binding to non-B-type DNA forms (*e.g.*, hemicatenated DNA loops, DNA minicircles, four-way junctions and cisplatin-modified DNA, K_d ~10^{-12}–10^{-7}, arranged in the descending affinity for the DNA structures, [43, 44, 48, 49]) is a characteristic feature of most, if not all, HMG-box proteins studied so far, regardless of their sequence specificity in binding to linear DNA.

The function of HMGB-type proteins as architectural factors is determined not only by their DNA-binding properties, but also by their ability to interact with a plethora of proteins. The binding partners of HMGB1 *in vitro* include transcription factors, DNA repair proteins, site-specific recombination proteins, silencing complexes, viral proteins ([2] and refs therein, [50–55]), and recently identified topoisomerase IIα (Štros et al., Nucleic Acid Res., in press) and a protein component of telomerase mTERT (Kunická et al., submitted) (Fig. 3). Depending on the type of HMGB protein, the interaction with other proteins may involve either the acidic C-tail (20–30 amino acid residues, mostly glutamic or aspartic acids) and/or the HMG-box. Unlike the DNA-binding region of the HMG-box, the protein-binding surface of the HMG-box is ill defined, and involves a short stretch of amino acids that interact with the P*XX*P*X*P motif of other proteins [54].

HMGB1 is engaged in many DNA transactions, such as DNA replication, repair, site-specific recombina-

tion, genomic integrity and transcription [2, 6, 7]. There is a compelling piece of evidence that normal development requires the proper expression of all types of HMG proteins, HMGA, HMGN, and HMGB [7]. Deregulation of expression of the HMG-box proteins, including the HMGB type, has profound consequences to cellular transcription, resulting in severe developmental defects and cancer ([7] and refs therein).

HMGB1 can exert its functions not only in the nucleus, but can also act as a signaling molecule (cytokine) *via* interaction with several surface molecules, including the receptor for advanced glycation end-products (RAGE), Toll-like receptors (TLR) 2 and 4, thrombomodulin and syndecan (reviewed in [56, 57]). The cytokine activities of HMGB1 become manifest when this protein translocates from the nucleus to the cytoplasm and into the extracellular environment. Re-localization of HMGB1 from the nucleus to the cytoplasm and its secretion are triggered by acetylation [6, 56, 57] and phosphorylation [58]. HMGB1, but not HMGB2, acts as a "necrotic marker" for the immune system to signal tissue damage or distress, and to activate responses of the organism towards healing. The extracellular function of HMGB1 is beyond the scope of this article and has been extensively reviewed elsewhere [56, 57].

HMGB-type proteins and genomic stability

Genomic instability can occur through a variety of mechanisms, including a defective response to DNA damage, and defects in DNA replication, or in chromosome segregation [59]. Loss of the yeast HMGB-type proteins, NHP6A and NHP6B (each containing a single HMG-box), results in increased genomic instability and hypersensitivity to DNA-damaging agents [60]. Similarly, HMGB1 deficiency in primary mouse embryonic fibroblasts (MEFs) (derived from *HMGB1*$^{-/-}$ mice, [37]) is accompanied by striking chromosomal instability, as evidenced by high levels of aneuploidy and spontaneous chromosome aberrations including chromosome breaks, and end-to-end chromosome fusions, dicentrics, rings, and Robertsonian-like fusions [60].

One mechanism for chromosome instability is through the dysfunction of telomeres, the specific nucleoprotein structures forming chromosome ends and protecting them from recombination and degradation. Replicative shortening can ultimately result in the uncapping of telomeres, with the loss of their protective function, and the unprotected end is further treated by the cell as a chromosome break. To counteract this "replicative shortening" cells can activate telomerase, a nucleoprotein complex with reverse-transcriptase activity, which can elongate 3' overhangs, while the complementary strand is filled in by conventional DNA polymerases [61]. In humans, telomerase is ubiquitously expressed only during the first weeks of embryogenesis, and is subsequently down-regulated in most cell types. Activation of telomerase is associated with about 90 % of cancers of various types [62]. Q-FISH analysis revealed numerous chromosomal aberrations in HMGB1-deficient MEFs, including end-to-end chromosome fusions lacking telomere signals (Fig. 4). Experiments aiming at understanding the role of HMGB1 in genomic stability identified HMGB1 as a telomerase-associated protein *in vitro* (Kunická et al., submitted). The possible involvement of HMGB1 in telomere structure and/or maintenance is supported by detection of markedly decreased telomerase activity, general telomere shortening and also by occasional telomere expansion (probably by the 'alternative lengthening of telomeres', ALT, pathway, [63]) in *HMGB1*$^{-/-}$ MEFs relative to the parental mouse cell line (Kunická et al., submitted). A dependence of telomerase activity on HMGB1 was also observed in human MCF-7 cells, where silencing of HMGB1 expression by specific short-hairpin RNA (shRNA) brought about a complete disappearance of telomerase activity relative to that in control cells (Kunická et al., submitted). Thus, the identification of HMGB1 as a factor involved in telomere organization may explain some aspects of previously observed chromosome instability in the HMGB1-defficient MEFs [60].

A suppression of the ALT pathway was recently reported by the SP100 protein *via* sequestration of the MRE11/RAD50/NBS1 complex (involved in DNA synthesis and recombination), which is present within the ALT-associated PML bodies together with telomeric DNA and telomere-binding proteins [64]. It would be interesting to find out whether an HMG-box containing a splice variant of SP100, SP100-HMG (which also colocalizes with PML bodies [20]) can suppress ALT. An unsolved question remains whether other HMG-box proteins, including HMGB2 and/or HMGB3, also participate in telomere maintenance and genomic stability. Although many DNA- or protein-protein binding activities of HMGB1–3 proteins are interchangeable *in vitro* [6], their cellular functions seem to be distinct as HMGB1$^{-/-}$ mice die 24 h after birth [37], and HMGB2$^{-/-}$or HMGB3$^{-/-}$ mice are viable [38–40]. Whatever the contribution of other HMGB-type proteins in maintaining genomic stability may be, the discovery of chromosomal abnormalities and telomere dysfunctions in HMGB1-deficient cells (Kunická et al., submitted

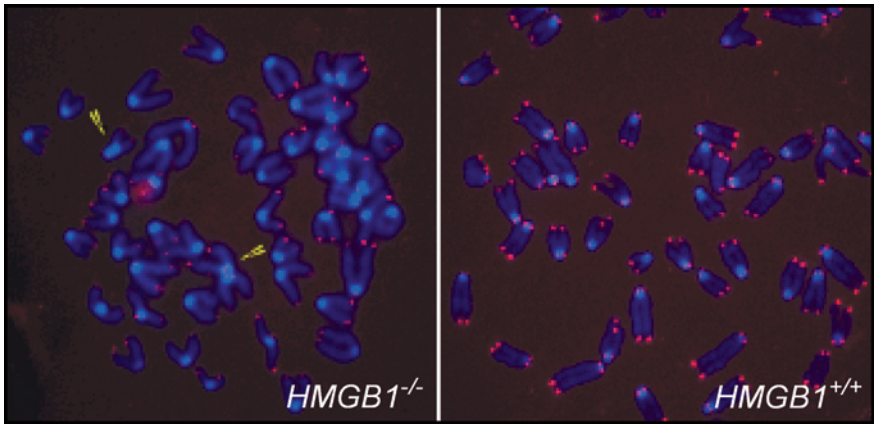

Figure 4. Increased chromosomal instability in HMGB1-deficient mouse embryonic fibroblasts (MEFs). Fluorescence *in situ* hybridization of metaphase spreads from mouse wild-type (*HMGB1*$^{+/+}$) and *HMGB1*$^{-/-}$ MEFs. Telomeres (red) were detected with Cy3-labeled PNA probe, and DNA stained with DAPI (blue). Arrows indicate chromosomal abnormalities in *HMGB1*$^{-/-}$ MEFs (Kunická et al., submitted). The parental and *HMGB1*$^{-/-}$ MEFs were kindly provided by, M. E. Bianchi; see also [60].

and [60] will undoubtedly initiate further research to clarify whether some aspects of chromosomal instabilities frequently found in malignant cells may also originate from aberrant expression of HMGB1 in tumors ([65] and refs therein).

The SRY and SOX proteins

In mammals, the process of sex determination commits the bipotential gonad to become either a testis (Sertoli cell) or an ovary (granulose cell) [66]. Differentiation of the precursor cells into Sertoli cells is determined by a single gene on the Y chromosome, encoding the SRY protein, which has been shown to be absolutely required for the development of male characteristics [66–69]. SRY contains a single evolutionarily highly conserved HMG-box playing a central role in the function of the protein, as nearly all missense mutations cause XY gonadal dysgenesis [70]. The exact molecular nature of how SRY initiate the program of gene expression to develop the bipotential embryonic gonad into a testis (rather than an ovary) is unknown. It involves sequence-specific DNA binding of its HMG-box to regulatory elements of a downstream target in Serotoli cell precursors, a testis-differentiation *SOX9* gene common to all vertebrates, encoding the SRY-related protein SOX9 [70].

The family of SOX proteins was discovered in 1990 as a group of proteins related to the mammalian testis-determining factor SRY [71]. The SOX proteins (~24–100 kDa) comprise nearly half of all human HMG-box proteins (Fig. 2), and the proteins are further divided into several distinct subgroups (Table 1). SOX proteins play important roles in a variety of developmental processes, particularly during organogenesis [30, 72], and only selected functions of individual SOX proteins are indicated in Table 1. *SOX* genes show diverse and dynamic patterns of expression throughout embryogenesis and in a variety of adult tissue types [30]. While *SRY* mutations result in sex reversal, mutations of several *SOX* genes result in severe developmental anomalies (such as a skeletal syndrome campomelic dysplasia associated with heterozygous mutations in SOX9). For more information, see the recent reviews on the SOX proteins [70, 73–75].

The SRY and SOX proteins bind DNA *via* a single sequence-specific HMG-box present in each of the proteins. Although all SOX/SRY proteins can recognize and bind a similar binding motif (A/TAACAA/T), different SOX proteins exhibit distinct preferences for the two nucleotides flanking the AACAAT motif [70]. The HMG-box seems to be interchangeable between the SOX proteins as revealed by HMG-box swapping experiments [76]. The HMG-box of SOX proteins interact not only with DNA, but it is also engaged in numerous protein-protein interactions, which are, however, primarily mediated by highly divergent non-HMG-box sequences ([73] and refs therein). In addition to the selectivity to slightly different DNA-binding sites within each of the target genes, the flexibility of the minor wing (*i.e.*, the extended N-terminal segment and helix III) enables SRY/SOX proteins to accommodate a broad range of sequence-specific DNA bend angles facilitating the binding of other DNA-binding proteins to adjacent binding sites [77]. Apart from the DNA-binding preferences of the individual SOX proteins, it is most likely the type of binding partner(s) in a particular tissue that determines the specificity of each member of the SOX family [75].

TCF/LEF-family and other HMG-box proteins in Wnt signaling

An example of co-operation of several distinct HMG-box proteins is Wnt signaling [78, 79]. Wnt signaling plays a crucial role in a variety of developmental

processes (such as specification of cell fate and polarity, and body axis formation), and its deregulation by mutations of several of its components [such as adenomatous polyposis coli tumor-suppressor protein (APC), axin, β-catenin and possibly also HMG-box proteins, see Fig. 5] results in malignant transformation (constitutive Wnt signaling). The essential event in Wnt signaling is the accumulation of β-catenin and the subsequent transcriptional activation by a complex of β-catenin and the sequence-specific HMG-box transcription factors TCF/LEF-1 [79]. An increase in β-catenin levels (through antagonization of the of the GSK3β kinase-APC-axin complex, see Fig. 5) increases the pool of nuclear β-catenin-TCF/LEF complexes, which, in turn, activate various Wnt/β-catenin target genes. In addition to TCF/LEF-1 proteins, Wnt/β-catenin signaling can also be modulated by other HMG-box proteins. These include two members of the SOX family, SOX17α/β and SOX3 from *Xenopus,* which repress Wnt signaling *via* physical interaction with β-catenin ([80] and refs therein). It is possible that human orthologues of *Xenopus* SOX17α/β and SOX3 may function in a similar way. Another HMG-box protein regulating Wnt signaling pathway is *C. elegans* SON-1 (sheath-to-neuron transformation-1) [81]. In the absence of Wnt signaling, TCF/LEF-1 transcription factors can associate with other binding factors, Groucho and CBP, acting as repressors [79]. While TCF/LEF-1 transcription factors are the best known HMG-box containing transcriptional activators of the Wnt pathway [12], two other human HMG-box proteins HBP1 (inhibitor of G1 and S phase progression; [32] and refs. therein) and HMG2L1 [80] function as transcriptional repressors of Wnt signaling (Fig. 5).

A constitutive Wnt signaling pathway is a pre-disposing event in many cancers of epithelial origin. Blockade of Wnt signaling and its target genes by HBP1 [32] and/or HMG2L1 [80] or other HMG-box proteins may provide a tumor-suppressive mechanism by blocking the oncogenic gene expression program. We can only speculate whether specific mutations within the HMG-box containing transcriptional repressors of Wnt signaling (loss-of-function) result in a constitutive Wnt signaling and cancer.

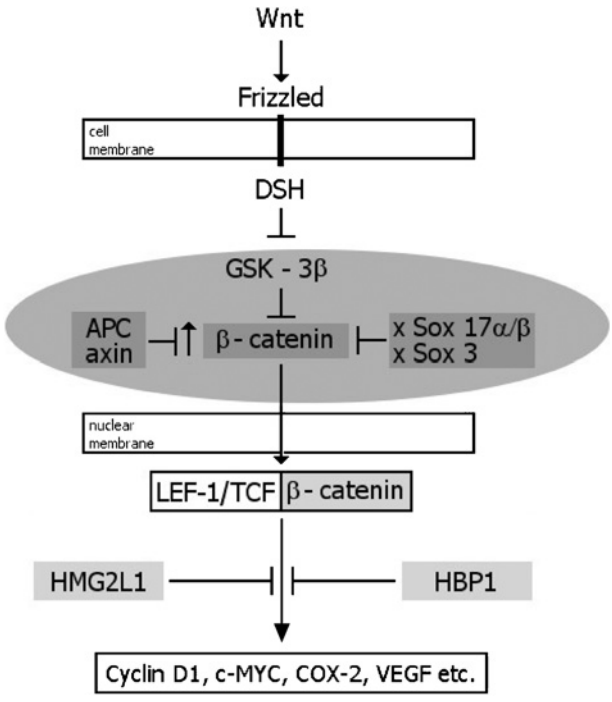

Figure 5. Wnt signaling and HMG-box proteins. Functioning of (i) inhibitors of accumulation of nuclear β-catenin SOX17α/β and SOX3, (ii) transcriptional activators TCF/LEF-1, and (iii) transcriptional repressors HBP1 and HMG2L1 is depicted. The picture is adapted from [32].

Chromatin remodeling

To make the chromatin-packed DNA accessible for gene transcription, DNA replication, repair or recombination, ATP-dependent chromatin remodeling multiprotein factors are required. Four different classes of nucleic acid-stimulating ATPase-containing remodeling complexes are recognized: SWI2/SNF2 (switching/sucrose non-fermenting-related), ISWI, Mi-2 and Ino80 [1, 82]. The human SWI/SNF remodelers can be subdivided further into two distinct, highly conserved subclasses, the BAF and PBAF complexes. Human cells contain two distinct Swi2/Snf2-like ATPase subunits, named hBRM (human Brahma) and BRG1 (Brahma-related gene 1). The SWI2/SNF2 family consists of a number of large multiprotein complexes containing between 8 and 10 subunits, always having a single catalytic subunit, BRM/SNF-A or BRG1/SNF-B, and several other variable BRG1-associated-factors (BAFs) that contribute to the enzymatic activity of the complex and facilitate the recruitment to sequence-specific transcription factors [82].

Moving of nucleosomes within chromatin is assured by ATP-dependent chromatin-remodeling factors. Compelling data exist suggesting an involvement of numerous HMG-box proteins in the modulation and maintenance of chromatin structure.

HMGB1

HMGB1 can increase the kinetics of nucleosome sliding (possibly by prebending the DNA at the edge of the nucleosome upon binding) catalyzed by the ATP-dependent ISWI-containing remodeling factors ACF and CHRAC [83]. HMGB1 and 2 are bound to

nucleosomal particles containing 180-bp DNA but not the core particles of 140-bp DNA, suggesting that these proteins bind to the linker segment of the nucleosomal DNA [84, 85], in close proximity to an exit/entry site on the core particle [86]. Involvement of HMGB1 in chromatin remodeling has recently been reviewed [87].

SSRP1

An attempt to identify factor(s) allowing RNA polymerase II to traverse nucleosomes (*i.e.*, elongation on chromatinized DNA templates) resulted in the isolation of "facilitates chromatin transcription" (FACT) as a factor that allows RNA polymerase II to traverse nucleosomes *in vitro* and *in vivo* by removing one H2A/H2B dimer [88–90]. In humans, FACT is a heterodimer consisting of Spt16 and an 80-kDa HMG-box protein structure-specific recognition protein 1 (SSRP1) [88]. SSRP1 plays also multiple roles in transcription regulation ([91] and refs therein). The C-terminal HMG-box of human SSRP1, like other HMG-box proteins [3], is able to recognize and bind specific DNA structures, such as DNA modified by the anticancer drug cisplatin [22] or cruciform DNA [92].

BRAF35

BRCA2-associated factor 35, BRAF35 (also named HMG20B) contains a non-sequence-specific HMG-box domain that can selectively bind four-way junctions ([93, 94] and refs therein). BRAF35 is present in a smaller complex (devoid of BRCA2 protein) with the capacity to deacetylate histones. The complex contains polypeptides reminiscent of the chromatin-remodeling complexes SWI/SNF and NuRD (nucleosome remodeling and deacetylating). BRAF35 and the closely related BAF57 are associated with BRG1, a human homologue of SWI2/SNF2 involved in ATP-dependent nucleosome disruption for activation of gene expression [95]. Its high similarity to the HMG-box-containing protein BAF57, the core subunit of the human BRG1-associated factors (BAF) complexes, suggests that BRAF35 may be involved in the regulation of gene expression through remodeling of chromatin structure.

PB1

The human PB1 (polybromo-1) protein was recently identified as a unique subunit of polybromo, BRG1-associated factors (PBAF) of a SWI/SNF chromatin-remodeling complex, which is required for localization of the PBAF complex at the kinetochores during mitosis [96, 97]. The 1634-amino acid PB1 protein contains six tandem bromodomains (BD), two bromo-adjacent homology domains (BAH), and an HMG-box. The HMG-box of PB1 differs from the known HMG-boxes by having a shortened third α-helix and an extended basic N terminus, which may bind to the minor groove [96]. The BDs bind acetylated histones, the BAH domains are protein-interaction modules, and the HMG-box has been shown to bind nucleosomal DNA. Thus, PB1 may serve as an important PBAF subunit coordinating events relating to chromatin remodeling complexes (targeting chromatin sites and recruiting specific proteins). Unlike the effect of HMGB1 in chromatin remodeling, which seems to be mediated solely by the HMG-boxes [83], the function of other HMG-box proteins (such as BRAF35, PB1 and SSRP1), apart from the HMG-boxes require also other domains participating mainly in protein-protein interactions.

Plant proteins containing HMG-box domains

Since the human genome has a gene number that is comparable to that of (some) plant genomes (both the human and the *Arabidopsis* genome contain ~25 000 protein coding genes), we intended to compare the entire complement of HMG-box proteins found in the human genome with that encoded in higher plant genomes. Therefore, we searched the Plant Chromatin Database (http://www.chromdb.org/) for protein sequences containing (putative) HMG-box domains. In our database survey we included the dicot species *Arabidopsis thaliana* and *Populus trichocarpa* (poplar), and the monocot species *Oryza sativa* (rice), whose genome sequences have been completely sequenced. Moreover, we have included the monocot species *Zea mays* (maize), because a variety of experimental studies of HMG-box proteins have been carried out analyzing maize proteins. The search revealed a total of 50 protein sequences that contain regions displaying striking amino acid sequence similarity to HMG-box domains. Therefore, it appears that compared to mammals, plant genomes (despite having a comparable total number of genes) encode a markedly smaller number of HMG-box proteins. The sequences of the plant HMG-box proteins (ranging from 13.6 to 72.5 kDa) were aligned and used to construct an amino acid sequence similarity tree that reveals the relationship between the protein sequences (Fig. 6). This analysis demonstrates that the plant HMG-box proteins can be subdivided into four groups: HMGB-type proteins, structure-specific recognition protein 1 (SSRP1), proteins containing three HMG-box domains (3xHMG-box), and proteins that contain both an AT-rich interaction domain (ARID) and an HMG-box domain (ARID/HMG). The latter two groups seem to be specific for plants. It appears that plant genomes do not encode HMG-box containing transcription factors such as SRY, LEF-1 and

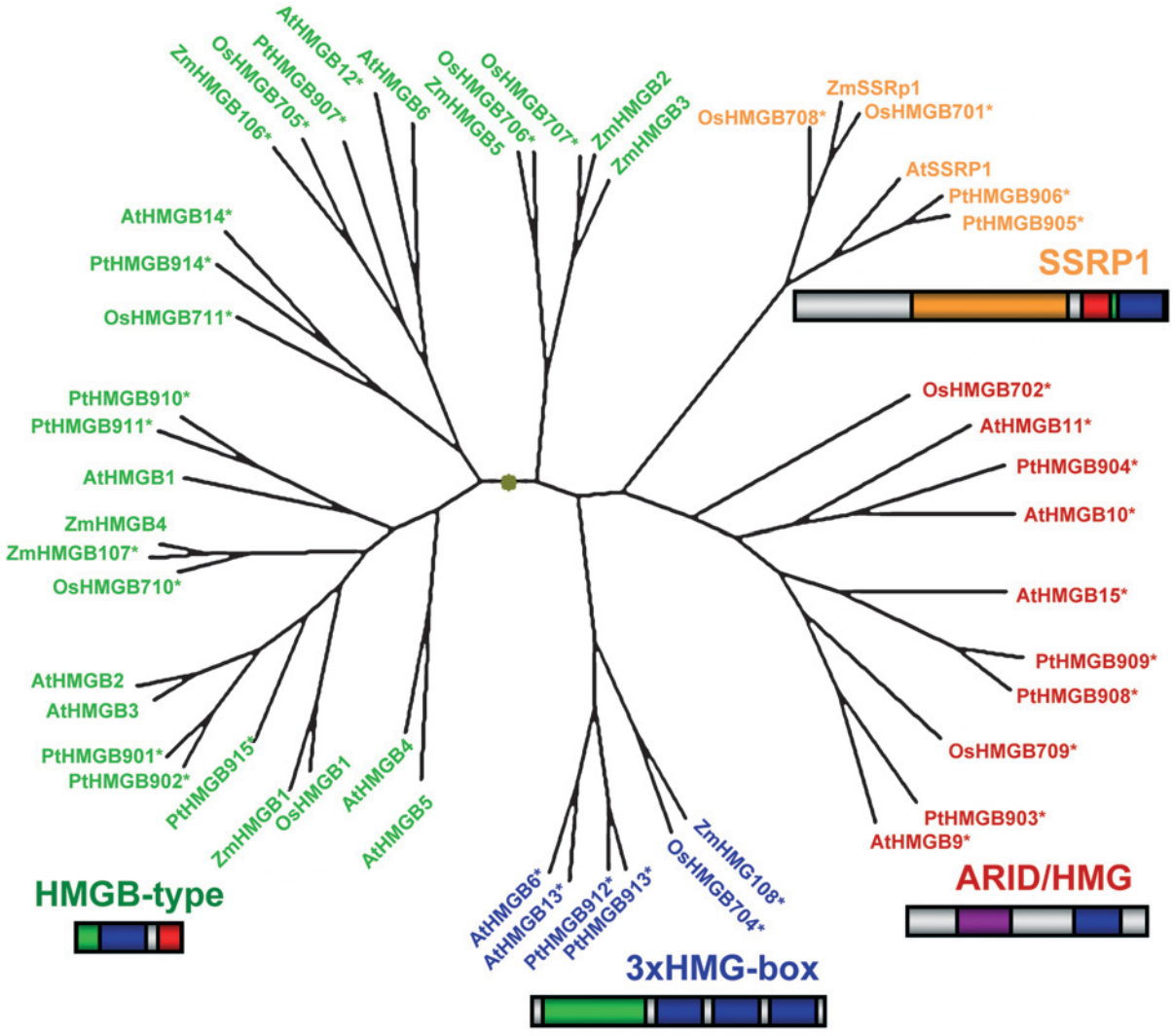

Figure 6. Overall structure and amino acid sequence similarity tree of plant HMG-box proteins. The amino acid sequences of HMG-box proteins from *Arabidopsis thaliana* (At), *Populus trichocarpa* (Pt), *Oryza sativa* (Os) and *Zea mays* (Zm) (Plant Chromatin Database, http://www.chromdb.org/) were aligned and a tree constructed as indicated in Figure 2. The published names of proteins were used if available, otherwise the name (indicated by asterisks) used in Plant Chromatin Database was used (*cf.* Table 2). The higher plant HMG-box proteins can be subdivided into four groups: HMGB-type proteins (green), 3xHMG-box proteins (blue), ARID/HMG proteins (red) and SSRP1 proteins (orange). Schematic representations of the overall structure of the four groups of HMG-box proteins that were identified in plants are included. Protein domains are indicated as follows: HMG-box domain (blue), basic region (green), acidic region (red), SSR domain of SSRP1 (orange) and ARID domain (violet).

the SOX family, which are highly divergent in mammals. The HMGB-type proteins represent by far the largest group of plant HMG-box proteins. In plants, HMGB-type proteins (13.6–27 kDa) are structurally relatively variable and around seven different HMGB-type proteins are found per species. They share a characteristic overall structure with a central HMG-box domain that is flanked by a basic N-terminal domain and an acidic C-terminal domain (Fig. 1a). The second largest subgroup of plant HMG-box proteins comprises proteins (35–51.9 kDa) that in addition to the HMG-box domain contain an ARID domain. The third group contains proteins (52.3–58 kDa) that have three HMG-box domains and an N-terminal region of unknown function. Monocot species contain a single gene coding for 3xHMG-box proteins, while the dicot species encode two proteins of this type. The fourth group represents the SSRP1 proteins (~70 kDa) that contain a central so-called SSR domain, followed by an acidic region and a C-terminal HMG-box domain. SSRP1 proteins are encoded by single genes in *Arabidopsis* and maize, while there are two genes in rice and poplar. We have summarized some characteristics of the plant HMG-box proteins encoded by the maize, rice, poplar and *Arabidopsis* genomes in Table 2.

The plant HMGB-type protein family

In contrast to mammalian HMGB-type proteins, which contain two HMG-box domains, the plant HMGB-type proteins are characterized by a single HMG-box domain that is flanked by a basic and an acidic domain (Fig. 1b). The HMG-box domain of the plant HMGB-type proteins displays a higher degree of similarity to the B domain of mammalian HMGB proteins rather than to the A domain [104]. Compared to mammals, plant HMGB-type proteins are structurally more variable, which is reflected by the very different sizes of the proteins. In *Arabidopsis*, for instance, HMGB-type proteins are found in the range of 14.2 kDa (HMGB5) and 27 kDa (HMGB6), which is essentially due to the variable N- and C-terminal domains, while the central HMG-box domain is relatively conserved. In plants, there are a greater number of different HMGB-type proteins per species (Table 2) than in mammals. Thus, the *Arabidopsis* genome encodes eight proteins that, based on their amino acid sequence, can be classified as HMGB-type proteins. However, it should be pointed out that for proteins that have been designated HMGB-type proteins based on their amino acid sequences, it has to be tested experimentally whether they share functional properties with *bona fide* HMGB proteins. A study of a putative *Arabidopsis* HMGB-type protein has revealed that despite striking amino acid sequence similarity to previously characterized HMGB proteins, AtHMGB12 (encoded by the AGI locus At5 g23405) has properties that are very different from those of typical HMGB proteins. The protein is mainly localized in the cytoplasm (rather than the nucleus) and it does not bind DNA [105].

The structural variability of plant HMGB-type proteins is also reflected by the fact that these proteins to some extent display different properties [42]. A recent study analyzing various *Arabidopsis HMGB* gene promoter-reporter gene fusions in transgenic plants has revealed that the *HMGB* genes are widely expressed in the plant, but there are also tissue- and developmental stage-specific differences in the expression patterns [106]. This suggests that the usual assumption that HMGB-type proteins are ubiquitously expressed needs to be modified. The expression of some *HMGB* genes appears to be regulated by an endogenous circadian rhythm [107]. Abiotic stress (drought, salinity, cold) can differentially modulate the expression of *HMGB* genes in *Arabidopsis*. Thus, the expression of *HMGB2*, *HMGB3* and *HMGB4* is up-regulated by cold stress, whereas the expression of *HMGB2* and *HMGB3* is down-regulated by drought or salt stress [108]. Different regions of *Arabidopsis* HMGB1 and HMGB5 are required for targeting the proteins to the cell nucleus, and they also differ somewhat in their nuclear mobility [109]. Maize HMGB-type proteins are differentially phosphorylated by protein kinase CK2 [110, 111]. The phosphorylation can enhance intramolecular interactions between the basic N-terminal and the acidic C-terminal domains reducing the affinity for linear DNA [112]. Moreover, the different HMGB-type proteins display differences in their DNA interactions (binding, bending) and their association with plant chromatin [42]. The efficiency of various maize HMGB-type proteins in promoting the formation of a nucleoprotein complex was tested using a site-specific recombination reaction as a model. In this assay, the HMGB-type proteins (depending on the type of DNA substrate) stimulated the β recombination reaction to different extents, indicating that the proteins differ in their ability to assist the assembly of nucleoprotein structures [113]. Similarly, the wheat and maize HMGB-type proteins have different potential to facilitate the binding of certain transcription factors to their target DNA sites [114, 115]. The functional interaction of the maize HMGB-type proteins with the transcription factor Dof2, is (as seen with other HMGB-transcription factor interactions) mediated by the HMG-box domain. The maize HMGB5 protein, for instance, is clearly most efficient in assisting the DNA-binding of Dof2. The HMG-box domains of HMGB1 and HMGB5 (in contrast to the full-length proteins) stimulate Dof2 DNA-binding to a similar extent, and it has been shown that the N- and/or C-terminal domains determine the efficiency of the functional HMGB protein interaction with Dof2 [116]. In summary, these findings support the view that various HMGB-type proteins may be adapted to serve partially different architectural functions in plant chromatin.

Currently, there are only few reports addressing the role of plant HMGB-type proteins in development or in response to changing environmental conditions. Ectopic expression of maize HMGB1 in tobacco seedlings results in reduced length of the primary root, whereas the shoot of the seedlings had wild-type appearance [117]. This "short-root" phenotype of the transgenic lines correlated with a decreased size of the cells in the cell division zone of the root. In these plants, the cell division rate rather than the cell elongation appears to be affected. A study addressing the correlation of *HMGB* gene expression and response to abiotic stress in *Arabidopsis* revealed interesting novel perspectives regarding HMGB protein function [108]. Overexpression of *HMGB2* and loss of function of *HMGB5* resulted in retarded seed germination and subsequent growth of the seedlings,

Table 2. Plant proteins containing HMG-box domain(s).

Protein[a]	Length (aa)	Mass (Da)	Locus[b]	Function[c]
HMGB-type				
AtHMGB1	178	20265	At3g51880	Structure-specific DNA binding, NL [99, 109]
AtHMGB2	144	15982	At1g20693	Structure-specific DNA binding [99]
AtHMGB3	141	15681	At1g20696	Structure-specific DNA binding [99]
AtHMGB4	138	15364	At2g17560	Structure-specific DNA binding [99]
AtHMGB5	125	14203	At4g35570	Structure-specific DNA binding, NL [99, 109]
AtHMGB6	241	26964	At5g23420	Structure-specific DNA binding, NL [100]
AtHMGB12*	149	16997	At5g23405	No structure-specific DNA binding+bending+NL [105]
AtHMGB14*	151	17481	At2g34450	Structure-specific DNA binding, no bending, NL [105]
PtHMGB901*	152	16531		
PtHMGB902*	152	16709		
PtHMGB907*	201	22544		
PtHMGB910*	176	19879		
PtHMGB911*	179	20033		
PtHMGB914*	165	19322		
PtHMGB915*	144	15916		
OsHMGB1	157	17100	Os06g51220	Structure-specific DNA binding, bending [102]
OsHMGB705*	202	22328	Os08g01100	
OsHMGB706*	145	15765	Os02g44930	
OsHMGB707*	131	13873	Os04g47690	
OsHMGB710*	146	16349	Os09g37910	
OsHMGB711*	133	15566	Os01g47600	
ZmHMGB1	157	17146		Structure-specific DNA binding, bending, β recomb., NL [98, 101, 103, 113]
ZmHMGB2	139	15316		Structure-specific DNA binding, bending, β recomb., NL [98, 101, 103, 113]
ZmHMGB3	138	15007		Structure-specific DNA binding, bending, β recomb., NL [98, 101, 103, 113]
ZmHMGB4	126	14104		Structure-specific DNA binding, bending, β recomb., NL [98, 101, 103, 113]
ZmHMGB5	123	13637		Stimulation of transcription factor Dof2, β recomb. [113, 114, 116]
ZmHMGB106*	212	23534		
ZmHMGB107	127	14215		
ARID/HMG				
AtHMGB9	338	38774	At1g76110	
AtHMGB10	319	36296	At3g13350	
AtHMGB11	337	38049	At1g55650	
AtHMGB15	448	50004	At1g04880	
PtHMGB903*	329	37740		
PtHMGB904*	316	36023		
PtHMGB908*	389	43825		
PtHMGB909*	467	51892		
OsHMGB702*	467	51630	Os02g27060	
OsHMGB709*	300	34993	Os09g37250	
3xHMG-box				
AtHMGB6*	456	53231	At4g23800	
AtHMGB13	446	52303	At4g11080	
PtHMGB912*	498	58004		
PtHMGB913*	480	56174		
OsHMGB704	504	56024	Os02g15810	
ZmHMGB108	487	55953		
SSRP1				
AtSSRP1	646	71646	At3g28730	NL, associates with transcribed genes [125]
PtHMGB905*	644	72498		
PtHMGB906*	610	68914		
OsHMGB701*	641	71334	Os01g08970	
OsHMGB708*	640	71048	Os05g08970	
ZmSSRP1	639	71418		Structure-specific DNA binding, bending, NL [122, 124]

[a] Published name of protein or * name used in Plant Chromatin Database (http://www.chromdb.org/) with sequences from *Arabidopsis thaliana* (At), *Zea mays* (Zm), *Oryza sativa* (Os) and *Populus trichocarpa* (Pt).
[b] Gene locus identifier, available for *Arabidopsis* (Arabidopsis Genome Initative, http://www.arabidopsis.org/) and rice (TIGR Rice Genome Annotation, http://www.tigr.org/tdb/e2k1/osa1/).
[c] Protein functions determined experimentally, including evidence for nuclear localization (NL) of the protein. References for the experimental work are given.

while altering the expression of *HMGB4* had no significant effect on these processes. In comparison to control plants, the expression of various germination-responsive genes was altered in *HMGB2*-overexpressing plants under salt stress. These results suggest that various HMGB proteins have different abilities to modulate germination and seedling growth in response to environmental stress [108].

Other plant proteins containing HMG-box domains

Proteins containing HMG-box and ARID domains

The second group of HMG-box proteins is characterized by the presence of both a C-terminal HMG-box domain and an N-terminal ARID domain (Fig. 6). Proteins with this combination of ARID/HMG-box domains appear to be specific for plants [118]. The ARID domain of ~100 amino acid residues was first identified in the murine Bright and the *Drosophila* Dead Ringer proteins and it occurs in proteins that are implicated in control of cell growth, differentiation and development [119]. The majority of tested ARID proteins appear to bind DNA without obvious sequence preference [120]. Currently, the function of only few proteins containing an ARID domain is known. Some have been suggested to act as transcription factors, while others are components of SWI/SNF-type ATP-dependent chromatin remodeling complexes [119]. The plant ARID/HMG proteins have not been examined experimentally. Therefore, the biological role of these plant-specific proteins remains to be clarified.

Proteins containing three HMG-box domains

The third group of HMG-box proteins comprises proteins with three HMG-box domains and an N-terminal region that displays no striking similarity to known protein domains (Fig. 6). These 3xHMG-box proteins appear to be specific for plants, but among plants they are relatively conserved. The *Arabidopsis* 3xHMG-box proteins are predicted to localize to plastids (http://www.arabidopsis.org/), but this requires experimental confirmation. Since this class of plant HMG-box proteins has not been analyzed experimentally, the cellular role of these proteins has still to be determined.

SSRP1

The SSRP1 is well characterized as a subunit of the FACT complex, which as a histone chaperone assists ATP-independent chromatin remodeling and facilitates transcript elongation of chromatin templates [90]. Human FACT activity is greatly stimulated by post-translational modification of histones, specifically by monoubiquitination of lysine 120 of human core histone H2B [121]. Plant SSRP1 has an overall structure closely related to that of mammalian SSRP1, although in the plant SSRP1 proteins, the domain C-terminal of the HMG-box domain is very short (Fig. 6). The HMG-box domain of maize SSRP1 mediates interaction with nucleosomes and non-sequence-specific binding to linear DNA, but it displays a preference for the structurally flexible dinucleotide step TG [122, 123]. Moreover, plant SSRP1 binds DNA structure selectively, and it can bend DNA as demonstrated by circularization assays [122, 124]. Similar to mammalian SSRP1, *Arabidopsis* SSRP1 associates with the 130-kDa protein Spt16 to form the FACT complex [125]. *Arabidopsis* FACT is found in the euchromatin of most cells, but not in the nuclei of terminally differentiated cells including cells of the root cap and trichoblasts. Consistent with a role in transcript elongation, FACT is detected by chromatin immunoprecipitation along the entire transcribed region of genes, but it is hardly found in intergenic and heterochromatic regions [125]. Various studies have indicated that SSRP1 also regulates gene transcription independently of Spt16 (*i.e.*, independently of FACT) [23].

Conclusions

The HMG-box domain represents a very versatile protein domain that mediates the DNA-binding of non-sequence-specific as well as of sequence-specific proteins. Typically, the HMG-box domain displays a high selectivity for certain DNA structures including minicircles and four-way junction DNA. Moreover, it is involved in different protein-protein interactions such as the functional interactions between HMGB-type proteins and various transcription factors. Accordingly, HMG-box domains are found in a wide variety of proteins. In humans, the most prominent group of HMG-box proteins are the SOX-type transcription factors, while no related proteins appear to exist in plants. In plants, the chromosomal HMGB-type proteins are the most divergent group of HMG-box proteins. Currently, the reason for the different evolution of the HMG-box motif in mammals and plants is unknown.

In line with their occurrence in a wide range of DNA-binding proteins, HMG-box proteins are involved in various nuclear (and extra-nuclear) functions including modulating chromatin structure and genomic stability, as well as contributing to the regulation of differentiation and development. In many cases, the HMG-box proteins serve an architectural role or act as chaperones that promote the efficient assembly of

higher-order regulatory nucleoprotein complexes. Identification of protein binding partners and the DNA target sites of the HMG-box proteins, and elucidation of the mechanisms that control of the activity of the HMG-box proteins *via* cellular signaling pathways will help to uncover how the members of this exciting protein family are integrated in the regulatory networks that determine the development of eukaryotic organisms.

Note added in proof

Interaction of HMGB1 with topoisomerase IIα indicated in Fig. 3 has recently been published: Štros M., Bačíková A., Polanská E., Štokrová J. and Strauss F. (2007) HMGB1 interacts with human topoisomerase IIα and stimulates its catalytic activity. *Nucleic Acid Res.*, in press.

Acknowledgements. Research in the laboratory of, K.D.G. has been supported by grants from the Danish Research Council and, D.L. is recipient of a DBRA stipend. Research in the laboratory of, M.Š. was supported by grants from the Grant Agency of the Czech Republic (204/05/2031) and the Grant Agency of the Academy of Sciences of the Czech Republic (IAA400040702). We also thank Terezie Keprtová for perfect secretarial assistance in preparation of the manuscript.

1 Längst, G. and Becker, P. B. (2004) Nucleosome remodeling: one mechanism, many phenomena? Biochim. Biophys. Acta 1677, 58–63.
2 Agresti, A. and Bianchi, M. E. (2003) HMGB proteins and gene expression. Curr. Opin. Genet. Dev. 13, 170–178.
3 Thomas, J. O. and Travers, A. A. (2001) HMG1 and 2, and related architectural DNA-binding proteins. Trends Biochem. Sci. 26, 167–174.
4 Bustin, M. (2001) Revised nomenclature for high mobility group (HMG) chromosomal proteins. Trends Biochem. Sci. 26, 152–153.
5 Jantzen, H. M., Admon, A., Bell, S. P. and Tjian, R. (1990) Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature 344, 830–836.
6 Bianchi, M. E. and Agresti, A. (2005) HMG proteins: dynamic players in gene regulation and differentiation. Curr. Opin. Genet. Dev. 15, 496 – 506.
7 Hock, R., Furusawa, T., Ueda, T. and Bustin, M. (2007) HMG chromosomal proteins in development and disease. Trends Cell Biol. 17, 72–79.
8 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.
9 Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C. and Norman, D. G. (1993) Solution structure of a DNA-binding domain from HMG1. Nucleic Acids Res. 21, 3427–3436.
10 Hardman, C. H., Broadhurst, R. W., A. R., C. Raine, A. R., C., 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.
11 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.
12 Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R. and Wright, P. E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. Nature 376, 791–795.
13 Murphy, E. C., Zhurkin, V. B., Louis, J. M., Cornilescu, G. and Clore, G. M. (2001) Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. J. Mol. Biol. 312, 481–499.
14 Masse, J. E., Wong, B., Yen, Y. M, Allain, F. H., Johnson, R. C. and Feigon, J. (2002) The, *S. cerevisiae* architectural HMGB protein NHP6A complexed with DNA: DNA and protein conformational changes upon binding. J. Mol. Biol. 18, 263–384.
15 Murphy, F. V., Sehy, J. V., Dow, L. K., Gao, Y. G. and Churchill M, E. (1999) Co-crystallization and preliminary crystallographic analysis of the high mobility group domain of HMG-D bound to DNA. Acta Crystallogr., D. Biol. Crystallogr. 55, 1594–1597.
16 Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O. and Lippard, S. J. (1999) Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. Nature 399, 708–712.
17 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.
18 Janke, C., Martin, D., Giraud-Panis, M. J., Decoville, M. and Locker, D. (2003) Drosophila DSP1 and rat HMGB1 have equivalent DNA binding properties and share a similar secondary fold. J. Biochem. (Tokyo) 133, 533–539.
19 Kamau, E., Bauerle, K. T. and Grove, A. (2004) The *Saccharomyces cerevisiae* high mobility group box protein HMO1 contains two functional DNA binding domains. J. Biol. Chem. 279, 55234–55240.
20 Guldner, H. H., Szostecki, C., Schroder, P., Matschl, U., Jensen, K., Luders, C., Will, H. and Sternsdorf, T. (1999) Splice variants of the nuclear dot-associated Sp100 protein contain homologies to HMG-1 and a human nuclear phosphoprotein-box motif. J. Cell Sci. 112, 733 – 747.
21 Yamada, M., Ohkawara, B., Ichimura, N., Hyodo-Miura, J., Urushiyama, S., Skirakabe, K. and Shibuya, H. (2003) Negative regulation of Wnt signalling by HMG2L1, a novel NLK-binding protein. Genes Cells 8, 677–684.
22 Bruhn, S. L., Pil, P. M., Essigmann, J. M. Housman, D. E. and Lippard, S. J. (1992) Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin. Proc. Natl. Acad. Sci. USA 89, 2307–2311.
23 Li, Y., Zeng, S. X., Landais, I. and Lu, H. (2007) Human SSRP1 has Spt16-dependent and -independent roles in gene transcription. J. Biol. Chem. 282, 6936–6945.
24 Kang, D., Kim, S. H. and Hamasaki, N. (2007) Mitochondrial transcription factor A (TFAM): Roles in maintenance of mtDNA and cellular functions. Mitochondrion 7, 39 – 44.
25 Bazett Jones, D. P., Leblanc, B., Herfort, M. and Moss, T. (1994) Short-range DNA looping by the *Xenopus* HMG-box transcription factor, xUBF. Science 264, 1134–1137.
26 Russell, J. and Zomerdijk, J. C. (2005) RNA-polymerase-I-directed rDNA transcription, life and works. Trends Biochem. Sci. 30, 87–96.
27 Sumoy, L., Carim, L., Escarceller, M., Nadal, M., Gratacos, M., Pujana, M. A., Estivill, X. and Peral, B. (2000) HMG20A and HMG20B map to human chromosomes 15q24 and 19p13.3 and constitute a distinct class of HMG-box genes with ubiquitous expression. Cytogenet. Cell. Genet. 88, 62–67.
28 Horii, A., Han, H. J., Sasaki, S., Shimada, M. and Nakamura, Y. (1994) Cloning, characterization and chromosomal assignment of the human genes homologous to yeast PMS1, a member of mismatch repair genes. Biochem. Biophys. Res. Commun. 204, 1257–1264.

29 O'Flaherty, E. and Kaye, J. (2003) TOX defines a conserved subfamily of HMG-box proteins. BMC Genomics 4, 13.
30 Pevny, L. H. and Lovell-Badge, R. (1997) Sox genes find their feet. Curr. Opin. Genet. Dev. 7, 338–344.
31 Lee, C. J., Chan, W. I., Cheung, M., Cheng, Y. C., Appleby, V. J., Orme, A. T. and Scotting, P. J. (2002) CIC, a member of a novel subfamily of the HMG-box superfamily, is transiently expressed in developing granule neurons. Mol. Brain Res. 106, 151–156.
32 Yee, A. S., Paulson, E. K., McDevitt, M. A., Rieger-Christ, K., Summerhayes, I., Berasi, S. P., Kim, J., Huang, C. Y. and Zhang, X. (2004) The HBP1 transcriptional repressor and the p38 MAP kinase: unlikely partners in G1 regulation and tumor suppression. Gene 7, 1–13.
33 Stec, I., van Ommen, G. J. and den Dunnen, J. T. (2001) WHSC1L1, on human chromosome 8p11.2, closely resembles WHSC1 and maps to a duplicated region shared with 4p16.3. Genomics 76, 5–8.
34 Sessa, L. and Bianchi, M. E (2007) The evolution of high mobility group box (HMGB) chromatin proteins in multicellular animals. Gene 387, 133–140.
35 Goodwin, G. H., Sanders, C. and Johns, E. W. (1973) A new group of chromatin-associated proteins with a high content acidic and basic amino acids. Eur. J. Biochem. 38, 14–19.
36 Vaccari, T., Beltrame, M., Ferrari, S. and Bianchi, M. E. (1998) Hmg4, a new member of the Hmg1/2 gene family. Genomics 49, 247–252.
37 Calogero, S., Grassi, F., Aguzzi, A., Voigtlander, T., Ferrier, P. Ferrari, S. and Bianchi, M. E. (1999) The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in new born mice. Nat. Genet. 22, 276–280.
38 Ronfani, L., Ferraguti, M., Croci, L., Ovitt, C. E., Schöler, H. R., Consalez, G. G. and Bianchi, M. E. (2001) Reduced fertility and spermatogenesis defects in mice lacking chromosomal protein Hmgb2. Development 128, 1265–1273.
39 Nemeth, M. J., Curtis, D. J., Kirby, M. R., Garrett-Beal, L. J., Seidel, N. E., Cline, A. P. and Bodine, D. M. (2003) Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation. Blood 102, 1298–1306.
40 Nemeth, M. J., Cline, A. P., Anderson, S. M., Garrett-Beal, L. J. and Bodine, D. M. (2005) Hmgb3 deficiency deregulates proliferation and differentiation of common lymphoid and myeloid progenitors. Blood 105, 627–634.
41 Štros, M., Štokrová, J. and Thomas, J. O. (1994) DNA looping by the HMG-box domains of HMG1 and modulation of DNA binding by the acidic C-terminal domain. Nucleic Acids Res. 22, 1044–1051.
42 Grasser, K. D., Launholt, D. and Grasser, M. (2007) High mobility group proteins of the plant HMGB family: dynamic chromatin modulators. Biochim. Biophys. Acta 1769, 346–357.
43 Pil, P. M., Chow, C. S. and Lippard, S. J. (1993) High-mobility-group 1 protein mediates DNA bending as determined by ring closures. Proc. Natl. Acad. Sci. USA 90, 9465–9469.
44 Paull, T. T., Haykinson, M. J. and Johnson, R. C. (1993) The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. Genes Dev. 7, 1521–1534.
45 Štros, M. (1998) DNA bending by the chromosomal protein HMGB1 and its high mobility group box domains. Effect of flanking sequences. J. Biol. Chem. 273, 10355–10361.
46 Grasser, K. D., Teo, S. H., Lee, K. B., Broadhurst, R. W., 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.
47 Štros, M. and Muselíková, E. (2000) 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. 275, 35699–35707.
48 Bianchi, M. E., Beltrame, M. and Paonessa, G. (1989) Specific recognition of cruciform DNA by nuclear protein HMG1. Science 243, 1056–1059.
49 Jaouen, S., de Koning, L., Gaillard, C., Muselíková-Polanská, E., Štros, M. and Strauss, F. (2005) Determinants of specific binding of HMGB1 protein to hemicatenated DNA loops. J. Mol. Biol. 353, 822–837.
50 Štros, M., Ozaki, T., Bačíková, A., Kageyama, H. and Nakagawara, A. (2002) HMGB1 and HMGB2 cell-specifically down-regulate the p53- and p73-dependent sequence-specific transactivation from the human Bax gene promoter. J. Biol. Chem. 277, 7157–7164.
51 Najima, Y., Yahagi, N., Takeuchi, Y., Matsuzaka, T., Sekiya, M., Nakagawa, Y., Amemiya-Kudo, M., Okazaki, H., Okazaki, S., Tamura, Y., Iizuka, Y., Ohashi, K. et al. (2005) High mobility group protein-B1 interacts with sterol regulatory element-binding proteins to enhance their DNA binding. J. Biol. Chem. 280, 27523–27532.
52 Yuan, F., Gu, L., Guo, S., Wang, C. and Li, G-M. (2004) Evidence for involvement of HMGB1 protein in human DNA mismatch repair. J. Biol. Chem. 279, 20935–20940.
53 Reddy, M. C., Christensen, J. and Vasquez, K. M. (2005) Interplay between human high mobility group protein 1 and replication protein A on psoralen-cross-linked DNA. Biochemistry 44, 4188–4195.
54 Dintilhac, A. and Bernues, J. (2002) HMGB1 interacts with many apparently unrelated proteins by recognizing short amino acid sequences. J. Biol. Chem. 277, 7021–7028.
55 Kamitani, W., Shoya, Y., Kobayashi, T., Watanabe, M., Lee, B. J., Zhang, G., Tomonaga, K. and Ikuta, K. (2001) Borna disease virus phosphoprotein binds a neurite outgrowth factor, amphoterin/HMG-1. J. Virol. 75, 8742–8751.
56 Lotze, M. T. and Tracey, K. J. (2005) High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. Nat. Rev. Immunol. 5, 331–342.
57 Raucci, A., Palumbo, R. and Bianchi, M. E. (2007) HMGB1: A signal of necrosis. Autoimmunity 40, 285–289.
58 Youn, J. H., And Shin, J. S. (2006) Nucleocytoplasmic shuttling of HMGB1 is regulated by phosphorylation that redirects it toward secretion. J. Immunol. 177, 7889–7897.
59 Tlsty, T. D. (1997) Genomic instability and its role in neoplasia, Curr. Top. Microbiol. Immunol. 221, 37–46.
60 Giavara, S., Kosmidou, E., Hande, M. P., Bianchi, M. E., Morgan, A., di Fagagna, F. A. and Jackson, S. P. (2005) Yeast Nhp6A/B and mammalian Hmgb1 facilitate the maintenance of genome stability. Curr. Biol. 15, 68–72.
61 Fajkus, J., Sýkorová, E. and Leicht, A. R. (2005) Telomeres in evolution and evolution of telomeres. Chromosome Res. 13, 469–479.
62 Shay, J. W. (1997) Telomerase in human development and cancer. J. Cell. Physiol. 173, 266–270.
63 Cerone, M. A., Londono-Vallejo, J. A. and Bacchetti, S. (2001) Telomere maintenance by telomerase and by recombination can coexist in human cells. Hum. Mol. Genet. 10, 1945–1952.
64 Jiang, W. Q., Zhong, Z. H., Henson, J. D., Neumann, A. A., Chang, A. C. and Reddel, R. R. (2005) Suppression of alternative lengthening of telomeres by Sp100-mediated sequestration of the MRE11/RAD50/NBS1 complex. Mol. Cell. Biol. 25, 2708–2721.
65 Völp, K., Brezniceanu, M. L., Bösser, S., Brabletz, T., Kirchner, T., Göttel, D., Joos, S. and Zörnig, M. (2006) Increased expression of high mobility group box 1 (HMGB1) is associated with an elevated level of the antiapoptotic c-IAP2 protein in human colon carcinomas. Gut 55, 234–242.
66 Brennan, J. and Capel, B. (2004) One tissue, two fates: molecular genetic events that underlie testis versus ovary development. Nat. Rev. Genet. 5, 509–521.
67 Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A. M., Lovell-Badge, R. and Goodfellow, P. N. (1990) A gene from the human sex-determining region encodes a protein with

homology to a conserved DNA-binding motif. Nature 19, 240–244.
68 Berta, P., Hawkins, J. R., Sinclair, A. H., Taylor, A., Griffiths, B. L., Goodfellow, P. N. and Fellous, M. (1990) Genetic evidence equating SRY and the testis-determining factor. Nature 348, 448–450.
69 Polanco, J. C. and Koopman, P. (2007) Sry and the hesitant beginnings of male development. Dev. Biol. 302, 13–24.
70 Harley, V. R., Clarkson, M. J. and Argentaro, A. (2003) The molecular action and regulation of the testis-determining factors, SRY (sex-determining region on the Y chromosome) and SOX9 [SRY-related high-mobility group (HMG) box 9]. Endocr. Rev. 24, 466–487.
71 Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature 346, 245–250.
72 Wegner, M. (1999) From head to toes: the multiple facets of Sox proteins. Nucleic Acids Res. 27, 1409–1420.
73 Wilson, M. and Koopman, P. (2002) Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. Curr. Opin. Genet. Dev. 12, 441–446.
74 Koopman, P., Schepers, G., Brenner, S. and Venkatesh, B. (2004) Origin and diversity of the SOX transcription factor gene family: genome-wide analysis in Fugu rubripes. Gene 328, 177–186.
75 Kanai, Y., Hiramatsu, R., Matoba, S. and Kidokoro, T. (2005) From SRY to SOX9: mammalian testis differentiation. J. Biochem. (Tokyo) 138, 13–19.
76 Kamachi, Y., Cheah, K. S., E. and Kondoh, H. (1999) Mechanism of regulatory target selection by the SOX high-mobility-group domain proteins as revealed by comparison of SOX1/2/3 and SOX9. Mol. Cell. Biol. 19**,** 107–120.
77 Weiss, M. A. (2001) Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition. Mol. Endocrinol. 15, 353–362.
78 Eastman, Q. and Grosschedl, R. (1999) Regulation of LEF-1/TCF transcription factors by Wnt and other signals. Curr. Opin. Cell Biol. 11, 233–340.
79 Hurlstone, A. and Clevers, H. (2002) T-cell factors: turn-ons and turn-offs. EMBO J. 21, 2303–2311.
80 Yamada, M., Ohkawara, B., Ichimura, N., Hyodo-Miura, J., Urushiyama, S., Skirakabe, K. and Shibuya, H. (2003) Negative regulation of Wnt signalling by HMG2L1, a novel NLK-binding protein. Genes Cells. 8, 677–684.
81 Jiang, L. I. and Sternberg, P. W. (1999) An HMG1-like protein facilitates Wnt signaling in *Caenorhabditis elegans*. Genes Dev. 13, 877–889.
82 Mohrmann, L. and Verrijzer, C. P. (2005) Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. Biochim. Biophys. Acta 1681, 59–73.
83 Bonaldi, T., Längst, G., Strohner, R., Becker, P. B. and Bianchi, M. E. (2002) The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. EMBO J. 21, 6865–6873.
84 Schröter, H. and Bode, J. (1982) The binding sites for large and small high-mobility-group (HMG) proteins. Studies on HMG–nucleosome interactions in vitro. Eur. J. Biochem. 127, 429–436.
85 Štros, M., Shick, V. V., Belyavsky, A. V. and Mirzabekov, A. D. (1985) Interaction of high mobility group proteins HMG 1 and HMG 2 with nucleosomes studied by gel electrophoresis. Mol. Biol. Rep. 10, 221–226.
86 Nightingale, K., Dimitrov, S., Reeves, R. and Wolffe, A. P. (1996) Evidence for a shared structural role for HMG1 and linker histones B4 and H1 in organizing chromatin. EMBO J. 15, 548–561.
87 Travers, A. A. (2003) Priming the nucleosome – a role for HMGB proteins? EMBO Rep. 4, 131–136.
88 Orphanides, G., Wu, W. H., Lane, W. S., Hampsey, M. and Reinberg, D. (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature 400, 284–288.
89 Svejstrup, J. Q. (2002) Chromatin elongation factors. Curr. Opin. Genet. Dev. 12, 156–161.
90 Reinberg, D. and Sims, R. J. (2006) de FACTo nucleosome dynamics. J. Biol. Chem. 281, 23297–23301.
91 Zeng, S. X., Dai, M. S., Keller, D. M. and Lu, H. (2002) SSRP1 functions as a co-activator of the transcriptional activator p63. EMBO J. 21, 5487**–**5497.
92 Gariglio, M., Ying, G. G., Hertel, L., Gaboli, M., Clerc, R. G. and Landolfo, S. (1997) The high-mobility group protein T160 binds to both linear and cruciform DNA and mediates DNA bending as determined by ring closure. Exp. Cell Res. 236, 472–481.
93 Marmorstein, L. Y., Kinev, A. V., Chan, G. K., Bochar, D. A., Beniya, H., Epstein, J. A., Yen, T. J. and Shiekhattar, R. (2001) A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression. Cell 104, 247–257.
94 Lee, Y. M. and Kim, W. (2003) Association of human kinesin superfamily protein member 4 with BRCA2-associated factor 35. Biochem. J. 374, 497–503.
95 Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M., Workman, J. L. and Crabtree, G. R. (1996) Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. EMBO J. 15, 5370–5382.
96 Goodwin, G. H. and Nicolas, R. H. (2001) The BAH domain, polybromo and the RSC chromatin remodelling complex. Gene 268, 1–7.
97 Chandrasekaran, R. and Thompson, M**.** (2006) Expression, purification and characterization of individual bromodomains from human Polybromo-1. Protein Expr. Purif. 50, 111–117.
98 Grasser, K. D., Wurz, A. and Feix, G. (1991) Isolation and characterisation of high-mobility-group proteins from maize. Planta 185, 350–355.
99 Stemmer, C., Ritt, C., Igloi, G. L., Grimm, R. and Grasser, K. D. (1997) Variability in *Arabidopsis thaliana* chromosomal high-mobility-group-1-like proteins. Eur. J. Biochem. 250, 646–652.
100 Grasser, K. D., Grill, S., Duroux, M., Launholt, D., Thomsen, M. S., Nielsen, B. V., Nielsen, H. K. and Merkle, T. (2004) HMGB6 from *Arabidopsis thaliana* specifies a novel type of plant chromosomal HMGB protein. Biochemistry 43, 1309–1314.
101 Ritt, C., Grimm, R., Fernández, S., Alonso, J. C. and Grasser, K. D. (1998) Four differently chromatin-associated maize HMG domain proteins modulate DNA structure and act as architectural elements in nucleoprotein complexes. Plant J. 14, 623–631.
102 Wu, Q., Zhang, W., Pwee, K. H. and Kumar, P. P. (2003) Rice HMGB1 protein recognizes DNA structures and bends DNA efficiently. Arch. Biochem. Biophys. 411, 105–111.
103 Wisniewski, J. R., Krohn, N. M., Heyduk, E., Grasser, K. D. and Heyduk, T. (1999) HMG1 proteins from evolutionary distant organisms distort B-DNA conformation in similar way. Biochim. Biophys. Acta 1447, 25–34.
104 Grasser, K. D (1995) Plant chromosomal high mobility group (HMG) proteins. Plant J. 7, 185–192.
105 Grasser, M., Lentz, A., Lichota, J., Merkle, T. and Grasser, K. D. (2006) The *Arabidopsis* genome encodes structurally and functionally diverse HMGB-type proteins. J. Mol. Biol. 358, 654–664.
106 Launholt, D., Grønlund, J. T., Nielsen, H. K. and Grasser, K. D. (2007) Overlapping expression patterns among the genes encoding *Arabidopsis* chromosomal high mobility group (HMG) proteins. FEBS Lett. 581, 1114–1118.
107 O'Neill, S. D. and Zheng, C. C. (1998) Abundance of mRNAs encoding HMG1/HMG2 class high-mobility-group DNA-

binding proteins are differentially regulated in cotyledons of *Pharbitis nil.* Plant Mol. Biol. 37, 235–241.
108 Kwak, K. J., Kim, J. Y., Kim, Y. O. and Kang, H. (2007) Characterization of transgenic *Arabidopsis* plants overexpressing high mobility group B proteins under high salinity, drought or cold stress. Plant Cell Physiol. 48, 221–231.
109 Launholt, D., Merkle, T., Houben, A., Schulz, A. and Grasser, K. D. (2006) *Arabidopsis* chromatin-associated HMGA and HMGB use different nuclear targeting signals and display highly dynamic localization within the nucleus. Plant Cell 18, 2904–2918.
110 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.
111 Stemmer, C., Leeming, D. J., Franssen, L., Grimm, R. and Grasser, K. D. (2003) Phosphorylation of maize and *Arabidopsis* HMGB proteins by protein kinase CK2α. Biochemistry 42, 3503–3508.
112 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.
113 Stemmer, C., Fernández, S., Lopez, G., Alonso, J. C. and Grasser, K. D. (2002) Plant chromosomal HMGB proteins efficiently promote the bacterial site-specific β-mediated recombination *in vitro* and *in vivo.* Biochemistry 41, 7763–7770.
114 Krohn, N. M., Yanagisawa, S. and Grasser, K. D. (2002) Specificity of the stimulatory interaction between chromosomal HMGB proteins and the transcription factor Dof2 and its negative regulation by protein kinase CK2-mediated phosphorylation. J. Biol. Chem. 277, 32438–32444.
115 Schultz, T. F., Spiker, S. and Quatrano, R. S. (1996) Histone H1 enhances the DNA binding activity of the transcription factor EmBP-1. J. Biol. Chem. 271, 25742–25745.
116 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.
117 Lichota, J., Ritt, C. and Grasser, K. D. (2004) Ectopic expression of the maize chromosomal HMGB1 protein causes defects in root development in tobacco seedlings. Biochem. Biophys. Res. Commun. 318, 317 – 322.
118 Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R., Creelman, R., Pilgrim, M. et al. (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. Science 290, 2105–2110.
119 Wilsker, D., Patsialou, A., Dallas, P. B. and Moran, E. (2002) ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. Cell Growth Differ. 13, 95–106.
120 Patsialou, A., Wilsker, D. and Moran, E. (2005) DNA-binding properties of ARID family proteins. Nucleic Acids Res. 33, 66–80.
121 Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatifard, A. and Reinberg, D. (2006) Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. Cell 125, 703–717.
122 Röttgers, K., Krohn, N. M., Lichota, J., Stemmer, C., Merkle, T. and Grasser, K. D. (2000) DNA-interactions and nuclear localisation of the chromosomal HMG domain protein SSRP1 from maize. Plant J. 23, 395–405.
123 Lichota, J. and Grasser, K. D. (2001) Differential chromatin association and nucleosome binding of the maize HMGA, HMGB and SSRP1 proteins. Biochemistry 40, 7860 – 7867.
124 Krohn, N. M., Stemmer, C., Fojan, P., Grimm, R. and Grasser, K. D. (2003) Protein kinase CK2 phosphorylates the HMG domain protein SSRP1 inducing the recognition of UV-damaged DNA. J. Biol. Chem. 278, 12710 – 12715.
125 Duroux, M., Houben, A., Ruzicka, K., Friml, J. and Grasser, K. D. (2004) The chromatin remodelling complex FACT associates with actively transcribed regions of the *Arabidopsis* genome. Plant J. 40, 660–671.