

# The HMGA proteins: A myriad of functions (Review)

ISABELLE CLEYNEN and WIM J.M. VAN DE VEN

Laboratory of Molecular Oncology, Department of Human Genetics, University of Leuven,  
Herestraat 49, B-3000 Leuven, Belgium

Received September 28, 2007; Accepted November 5, 2007

**Abstract.** The 'high mobility group' HMGA protein family consists of four members: HMGA1a, HMGA1b and HMGA1c, which result from translation of alternative spliced forms of one gene and HMGA2, which is encoded for by another gene. HMGA proteins are characterized by three DNA-binding domains, called AT-hooks, and an acidic carboxy-terminal tail. HMGA proteins are architectural transcription factors that both positively and negatively regulate the transcription of a variety of genes. They do not display direct transcriptional activation capacity, but regulate gene expression by changing the DNA conformation by binding to AT-rich regions in the DNA and/or direct interaction with several transcription factors. In this way, they influence a diverse array of normal biological processes including cell growth, proliferation, differentiation and death. Both HMGA1 and HMGA2 are hardly detectable in normal adult tissue but are abundantly and ubiquitously expressed during embryonic development. In malignant epithelial tumors as well as in leukemia, however, expression of HMGA1 is again strongly elevated to embryonic levels thus leading to ectopic expression of (fetal) target genes. HMGA2 overexpression also has a causal role in inducing neoplasia. Besides overexpression of full length HMGA proteins in different tumors, the *HMGA* genes are often involved in chromosomal rearrangements. Such translocations are mostly detected in benign tumors of mesenchymal origin and are believed to be one of the most common chromosomal rearrangements in human neoplasia. To provide clarity in the abundance of articles on this topic, this review gives a general overview of the nuclear functions and regulation of the *HMGA* genes and corresponding proteins.

## Contents

1. Introduction
2. The high mobility group of chromosomal proteins

3. The *HMGA* genes and corresponding proteins
4. The HMGA proteins: mechanism of action
5. Cellular functions of the HMGA proteins
6. Regulation of *HMGA* expression
7. Conclusion

## 1. Introduction

Life depends on the ability of a cell to store, retrieve, process and translate the genetic instructions required to make and maintain a living organism. This hereditary information is passed on from a cell to its daughter cells at cell division, and from generation to generation of organisms through the reproductive cells. Large amounts of DNA are required to encode all the information needed to generate the instructions for the development of a multicellular organism. This information needs to be stored within the eukaryotic cell and handled appropriately. The latter is achieved by packaging the enormously long double-stranded DNA molecules into chromosomes, which not only fit readily inside the nucleus but can be easily apportioned between the two daughter cells at each cell division. The complex task of packaging is accomplished by specialized proteins that bind to and fold the DNA, generating a series of coils and loops that provide increasingly higher levels of organization and that prevent DNA from becoming an unmanageable tangle. Amazingly, the DNA is compacted in such an orderly fashion so that the genes contained in the DNA molecules are available when they need to be transcribed or replicated. The chromatin fiber thus not only serves to package the DNA into the nucleus, but also provides a means to control the accessibility of specific sequences to regulatory factors and to potentiate interactions between distant regulatory elements. Most of the cellular processes involving DNA thus have to be considered in the context of chromatin (adapted from ref. 1).

## 2. The high mobility group of chromosomal proteins

In eukaryotic chromatin, proteins associate with DNA and condense the molecule into the nucleosomal filament made of the succession of nucleosome core particles linked together by DNA segments. Precise interactions between proteins and DNA in chromatin facilitate the orderly progression of complex processes such as transcription, replication, recombination, and repair. Most of the studies on the structure and function of chromatin have focused on interactions occurring between

---

*Correspondence to:* Dr Wim Van de Ven, Department of Human Genetics, University of Leuven, Herestraat 49/bus 602, B-3000 Leuven, Belgium  
E-mail: wim.vandeven@med.kuleuven.be

*Key words:* HMGA protein, DNA-binding

histones (around which the DNA is wrapped and which are constitutive components of the chromatin) and DNA. It is now clear that besides histones there is a second class of proteins, which are either part of, or associated with, the chromatin fiber. These proteins are referred to as non-histone chromosomal proteins and constitute a heterogeneous class of molecules that play an important role in the structure and function of chromosomes and in the regulation of gene activity.

The high mobility group (HMG) proteins are among the largest and best characterized group of non-histone chromosomal proteins (reviewed in refs. 2-4). They are defined as nuclear proteins with a high electrophoretic mobility in polyacrylamide gels, can be extracted from nuclei or chromatin with 0.35 M NaCl, and are soluble in 5% perchloric or trichloroacetic acid. Typically, they have a high content of charged amino acids and a molecular mass <30 kDa. The HMG protein family has been classified into three subfamilies: HMGA, HMGB and HMGN, previously known as HMGI/Y, HMG1/2, and HMG14/17 respectively (5). Each of the subfamilies has a unique protein signature and a characteristic functional sequence motif: the 'AT-hook' for the HMGA family, the 'HMG-box' for the HMGB family, and the 'nucleosomal binding domain' for the HMGN family. Via their respective functional motifs, HMG proteins are able to bind to specific structures in DNA or chromatin in a sequence-independent fashion, since they recognize structure rather than a particular nucleotide sequence. HMG proteins function as architectural elements that modify the structure of their binding partners to generate a conformation that facilitates and enhances various DNA-dependent activities such as transcription, replication, recombination, and repair.

The functional motifs characteristic to these canonical HMG proteins have also been identified in other, mostly nuclear, proteins which are called HMG-motif proteins. It is important to distinguish the HMG-motif proteins from the archetypal HMG proteins. The archetypal HMG proteins are ubiquitous in all the cells of higher eukaryotes, are relatively abundant and bind to the DNA in a sequence-independent fashion (also see above). On the contrary, the HMG-motif proteins are cell-type-specific, not abundant and bind to the DNA in a sequence-specific fashion (3). This review concentrates on the archetypal HMG protein family, more specifically on members of the HMGA family.

### 3. The HMGA genes and corresponding proteins

Over thirty years ago, the high mobility group proteins were discovered as abundant heterogeneous non-histone components of chromatin (6). In 1983, two new high-mobility group-like proteins were found in HeLa cells (7). Later, these small DNA-binding proteins were categorized as HMGA proteins and different scientific groups became interested in them.

*Two genes and four proteins.* The mammalian HMGA family consists of two functional members, HMGA1 and HMGA2 (Fig. 1). Chromosome mapping studies have located the HMGA1 gene on human chromosome 6 (6p21) and on mouse chromosome 17. The human HMGA1 gene contains eight exons, which are distributed over a region of about 10 kb (8),

whereas the mouse *hmga1* orthologue contains six exons spanning about 7 kb (9,10). On the other hand, the human HMGA2 gene is located at chromosomal band 12q14-15 and contains at least five exons dispersed over a genomic region of  $\geq 160$  kb. The mouse *hmga2* gene locates at the pygmy locus on chromosome 10 and contains five exons spanning more than 110 kb. Both human and mouse HMGA2/*hmga2* genes are much larger than their HMGA1/*hmga1* counterparts, mainly because of longer 5' and 3' untranslated regions and because of the extremely long third intron of HMGA2/*hmga2*.

Alternative splicing of the HMGA1 gene transcript gives rise to three mRNAs encoding HMGA1a (107 amino acids, 11.7 kDa), HMGA1b (96 amino acids, 10.6 kDa) and the more recently identified HMGA1c (179 amino acids, 19.7 kDa) (11). The HMGA1a and HMGA1b isoforms differ by only 11 amino acids present in HMGA1a but not in HMGA1b and are encoded by the most abundant splice variants of the HMGA1 gene. HMGA1c is produced from the HMGA1 gene by alternative splicing using noncanonical splice donor and acceptor sites. This alternative splicing results in a frameshift so that the two proteins are identical in their first 65 amino acids but differ thereafter.

In normal cells, transcripts from the HMGA2 gene code primarily for the full-length HMGA2 protein (109 amino acids, 12 kDa). HMGA proteins occur widespread in nature and homologues for the mammalian HMGA proteins have been found in yeast, insects, plants and birds, as well as in all mammalian species examined (reviewed in refs. 2 and 4).

*HMGA proteins and their unusual structural characteristics.* Amino acid alignment of the four HMGA proteins (Fig. 2) shows that the HMGA2 protein has a high amino acid sequence homology ( $\pm 55\%$ ) with the HMGA1a and HMGA1b proteins, including the presence of three conserved DNA-binding domains, also called AT-hooks (see below). At their C-termini, the HMGA proteins, except for HMGA1c, have a high percentage of negatively charged acidic residues. For both genes, the three independent DNA-binding domains are each located on separate exons, as are the regions coding for the acidic C-terminal domains.

The AT-hook motif is a positively charged stretch of 9 amino acids containing the invariant repeat Arg-Gly-Arg-Pro (R-G-R-P), flanked by other positively charged residues [usually Arg (R) and Lys (K); Fig. 2]. With these AT-hooks, the HMGA proteins recognize and bind to AT-rich sequences in the minor groove of B-form DNA. As stated above, they recognize structure, rather than a particular nucleotide sequence. Free in solution, AT-hooks possess little, if any secondary structure. Upon binding to DNA, the AT-hook undergoes a disordered-to-ordered conformational change. A single DNA-binding domain preferentially binds to stretches between four and six base pairs, or approximately half a turn of the double helix (4). Simultaneous binding of two or more AT-hooks to different binding sites in its DNA substrate results in an increase in the strength of interactions of HMGA proteins with DNA (12). Depending on the number and spacing of AT-rich binding sites in DNA, HMGA proteins can influence the conformation of bound DNA substrates in different ways, and consequently, affect numerous biological effects.

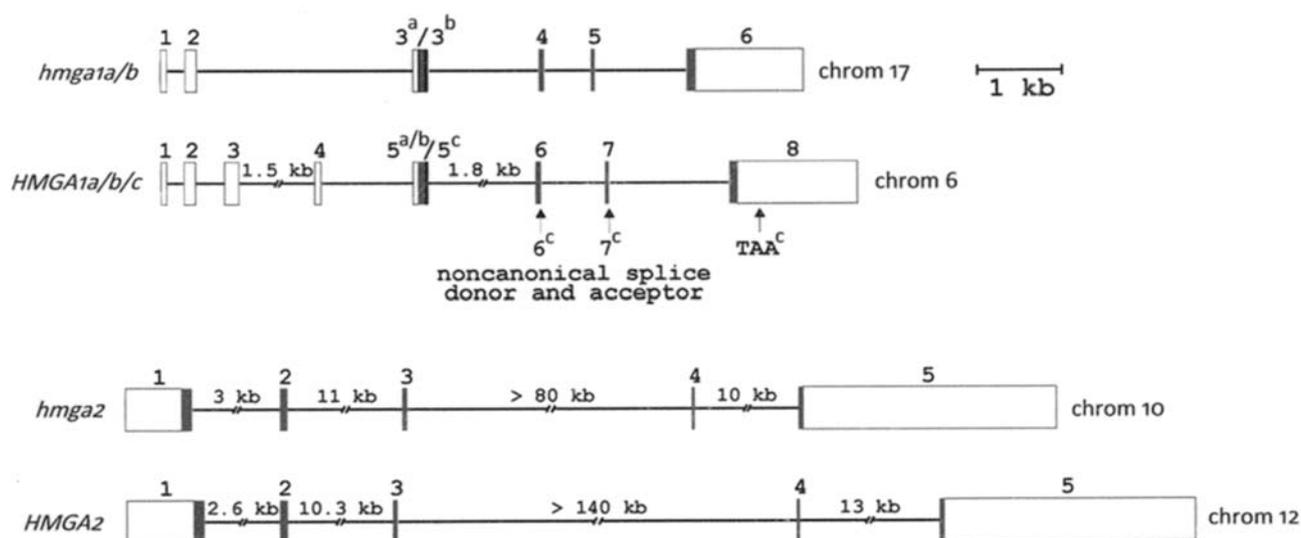


Figure 1. Schematic representation of the genomic organization of the mouse *hmga* and human *HMGA* genes. Introns and genomic sequences (lines) and exons (numbered boxes); coding sequences (dark grey boxes); 5' and 3' untranslated regions (open boxes); 33 nucleotides missing due to alternative splicing of exons 3 and 5 in mouse and human *HMGA1b* as compared to *HMGA1a* respectively (black boxes). The usage of non-canonical splice donor and acceptor sites of the human *HMGA1* gene results in the *HMGA1c* transcript that has a 3' shortened exon 6 (6<sup>c</sup>) and a 5' shortened exon 7 (7<sup>c</sup>); Taa<sup>c</sup>, translational termination codon of the *HMGA1c* transcript (arrows); chrom, chromosome. All sequences are drawn to scale, except when indicated (adapted from ref. 169).



Figure 2. Amino acid alignment of human HMGA proteins. AT-hooks (bold); invariant repeat R-G-R-P (grey box); C-terminal acid tail (red box; italics). The region that is deleted in *HMGA1b* as compared to *HMGA1a* (asterisks). The amino acid residues of *HMGA1c* that differ from *HMGA1a* (underlined) (adapted from ref. 169).

The function of the acidic tail is still not understood. It is believed that the C-terminal tail may be important in modulating protein-protein interactions (13,14) and could be involved in enhancing transcription factor activity (13). This latter role is, however, quite controversial, since there also exist examples in which removal of the acidic tail has no influence on enhancing activation of genuine transcription factors (15).

Interestingly, HMGA1 proteins also have the capacity to bind to certain types of structures formed by non AT-rich DNA sequences. For example, *in vitro*, the HMGA1 protein as well as its DNA-binding domain(s) alone bind to synthetic four-way junction structures, to non-B-form structures in supercoiled plasmids, and to distorted regions of DNA found on isolated nucleosome core particles (16-18).

In addition, HMGA proteins participate in protein-protein interactions and induce structural changes in the bound

protein substrates. These interactions are a second way in which the HMGA proteins participate in the regulation of expression of many different genes as will be explained below.

In conclusion, by binding to DNA and proteins, HMGA proteins induce structural changes in the substrates with which they interact. In this way, they act as architectural transcription factors that can influence the expression of many different genes. It should be emphasized again that they do not have intrinsic transcriptional activation capacity (19).

#### 4. The HMGA proteins: mechanism of action

HMGA proteins participate in a wide variety of nuclear processes ranging from chromosome and chromatin dynamics to acting as architectural transcription factors that regulate

the expression of numerous genes. Via protein-DNA and protein-protein interactions, they influence a diverse array of normal biological processes including cell growth, proliferation, differentiation and death. The mechanisms by which HMGA proteins exert their diverse biological functions will be further discussed.

#### *HMGA proteins and chromosome dynamics*

*Chromosome condensation.* The state of condensation of chromosomes varies dramatically during the various stages of the cell cycle. During mitosis, the chromosomes replicate, condense and are subsequently distributed to the two daughter nuclei. The condensed state of the mitotic chromosomes is important in allowing the duplicated chromosomes to be separated during cell division. In interphase nuclei, the chromosomes are in a more extended state, which allows access to the DNA for the enzymes and other proteins required for gene expression and DNA replication. The eukaryotic cell thus has the ability to vary the packaging of its DNA and uses this during progression through the cell cycle as a mechanism to facilitate the required processes at the various time-points.

It has to be noted that the interior of the nucleus is not a random jumble of its DNA, RNA and protein components. Each chromosome occupies a particular region of the nucleus so that different chromosomes do not become entangled with each other. This organization involves, at least in part, attachments of parts of the chromosomes to sites on the nuclear matrix, called matrix/scaffold associated regions (MARs/SARs). Functionally important are, in this context, the highly AT-rich sequences which have high affinity for the nuclear matrix and accordingly organize genomic DNA into topologically distinct loop domains that are important in transcription. In this regard, SARs often localize in close proximity to gene regulatory sequences, including enhancers (20).

HMGA1 proteins have been shown to be localized at the AT-rich G/Q- and C-bands as well as to the SARs of human and mouse metaphase chromosomes (21-24). These localizations suggest that the HMGA proteins are actively involved in the dynamic changes in chromatin structure that occur during chromosome condensation in the cell division cycle. Indeed, treatment of cells with distamycin - a drug which displaces histone H1 and HMGA1 from SARs and other AT-rich sequences - causes marked decondensation of centromeric heterochromatin and substantial metaphase chromosome elongation, indicating that both histone H1 and HMGA1 are necessary for chromosome condensation (25,26). However, it is also in such metaphase chromosomes that both the histone H1 and HMGA1 proteins are maximally phosphorylated by the cell cycle-dependent cdc2 kinase (see below), and most loosely bound to the substrate DNA. It is, therefore, possible that this hyperphosphorylation results in an alteration of binding affinity, hereby allowing association of histone H1 and HMGA1 with other chromatin proteins that are causally involved in DNA compaction (27).

It is well established that the chromatin in an interphase chromosome is not in the same packaging state throughout the whole chromosome. Some regions are more unfolded and extended than others. In general, chromosome regions that

are being transcribed into RNA are more extended, while those that are transcriptionally silent are more compact. The most highly condensed form of interphase chromatin is called 'heterochromatin'. It is typically concentrated around the centromere region and at the ends of the chromosomes (telomeres). Like mitotic chromatin, heterochromatin is transcriptionally inactive. Chromatin which can occur in a (variety of) more extended state(s) is called 'euchromatin'. Part of this euchromatin is in a state in which it is either actively being transcribed or easily available for transcription.

The HMGA1 proteins were first reported to be associated with the centromeric heterochromatin and have been proposed, based on their binding to discrete sites in African green monkey  $\alpha$ -satellite repeats, to play a role in nucleosomal phasing (28). It has also been shown that the HMGA1 proteins bind to multiple sites in mouse satellite DNA. Studies using the AT-binding drugs Hoechst 33258 and distamycin A show that these drugs cause an incomplete condensation of centromeric chromatin by competing with the HMGA1 proteins for binding to satellite DNA (26).

Together, these data point towards an important role for HMGA1 proteins in chromatin dynamics during the cell cycle as well as for regulation of transcription on the level of chromatin condensation.

*Inhibition of nucleotide excision repair.* The amount of information stored in a cell's DNA is enormous. Each cell not only contains an elaborate machinery for accurately copying this stored information, but also specialized enzymes for repairing DNA, when it becomes damaged. Mutations in DNA may affect the information it encodes. Occasionally, this can benefit the organism in which the mutation occurs. In fact, accumulation of DNA changes over millions of years is generally assumed to constitute the basis for the genetic variation among species and indeed is considered the basis of evolution. It also produces the small variations within the same species that we can easily see in humans and other animals. Mutations may, however, be detrimental, and, in humans reflect thousands of inherited diseases. Mutations that arise in the cells of the body throughout the lifetime of an individual may also cause disease, most notably the many types of cancer.

One example of DNA damaging agents is UV radiation. UV radiation produces stable DNA photoproducts by covalent linkage between adjacent bases. The most abundant forms of these photoproducts are cyclobutane pyrimidine dimers (CPD). CPDs alter DNA structure, causing severe bending of the helical axis. They can obstruct progression of DNA polymerase and can inhibit transcription by blocking RNA polymerase or by interrupting formation of the transcription initiation complex. If these CPDs are left unrepaired they can cause mutations and lead to for example skin cancer (29). It has been shown by Adair *et al* that HMGA1 overexpression in human cells inhibits their ability to survive exposure to relatively high doses of UV light and that it also significantly reduces their ability for general genomic nucleotide excision repair. One way in which HMGA1 can hamper nucleotide excision repair is by physically inhibiting access of DNA repair-associated proteins to CPD lesions, by directly binding to the UV-induced lesions (30,31). Likewise, HMGA1 proteins

might negatively regulate transcription of certain genes that code for repair factors, as is shown for the nucleotide excision repair factors ERCC1 (32), and Xeroderma Pigmentosum Group A (XPA) (33). The decreased repair activity could also be a result of direct interactions between HMGA1 proteins and repair factors, thereby inhibiting their function.

#### *Regulation of gene transcription in vivo*

**Chromatin structure and gene regulation.** As mentioned above, chromatin structure plays a dominant role in the regulation of gene transcription in eukaryotic cells. Genes which are being expressed seem to be in a more extended form of chromatin, although the DNA is still packaged into nucleosomes. Since nucleosomes are placed along DNA at regular intervals with little apparent sequence specificity, they are likely to occur over promoter regions. In that case, they can inhibit initiation of transcription because they prevent the general transcription factors or RNA polymerase from assembling on the DNA. Also other inhibitory chromatin proteins (e.g. histone H1) can be associated with critical regulatory regions of gene promoters or enhancers. When transcription of a gene is activated, such nucleosomes and other inhibitory proteins are displaced from the DNA. One way of achieving this displacement is by destabilizing the intermolecular binding of histones and other regulatory protein complexes in the vicinity of a promoter by biochemically modifying them. Another way is to have specialized proteins which can displace histones and other regulatory protein complexes from promoters and consequently clear the way for the general transcription factors to assemble. One class of these specialized proteins are the HMGA proteins. HMGA proteins act as regulators of gene transcription by controlling the global structure of large loops or domains of chromatin. Preferential binding of HMGA proteins to the SARs at the base of repressed chromatin loops and concomitant displacement/exclusion of histone H1 from these sequences leads to a local opening of chromatin and initiation of the gene activation process. This initial binding is followed by a cooperative propagation of HMGA binding throughout an entire chromatin domain and the establishment of an accessible structure that is permissive for gene transcription. In this model, the HMGA proteins act as 'anti-repressor' molecules that displace inhibitory proteins from SAR sequences (24). An example is the proximal promoter region of the human *IL-2R $\alpha$*  gene, where stably positioned nucleosomes that occlude important transcription factor binding sites are present on the PRRII regulatory element. The HMGA proteins are proposed to be involved in the initial remodelling of the inhibitory nucleosomes and the subsequent formation of an enhanceosome on this element [(34); also see below].

In this context, it has also been proposed that recruitment of histone acetyl transferases (HATs) by sequence-specific transcription factors, like the retinoic acid receptors (RARs) and retinoid X receptors (RXRs), leads to destabilization of nucleosomal cores by acetylation of core histones. RARs and RXRs are nuclear receptors which induce gene expression in a ligand-dependent manner through RA-responsive elements (RAREs) present in the promoter regions of responsive genes. However, before HATs can acetylate the core histones, the inhibitory linker histone H1 should first be removed. Since

HATs do not acetylate the histone H1 in its DNA binding tail, removal of histone H1 has to be achieved in another way. Nagpal *et al* showed that RARs are capable of recruiting the HMGA1 proteins, which can then displace histone H1. This results in a conversion from the silent 30-nm chromatin fiber to a relatively open 10-nm fiber. In a second step, HAT activities are then recruited, leading to a destabilization of the nucleosomal core and active transcription (11). These scientists were also the first to describe the existence of a third HMGA1 protein, namely HMGA1c (see above).

**Long-range chromatin interactions.** Nearly all eukaryotic promoters require activator proteins to aid the assembly of the general transcription factors and RNA polymerase. The DNA sites to which the eukaryotic gene activators bind are called 'enhancers'. These activator proteins can be bound thousands of nucleotide pairs away from the promoter. To allow the activator proteins bound to the enhancer to come into contact with either RNA polymerase or with one of the general transcription factors bound to the promoter, the DNA between the enhancer and the promoter has to loop out. The HMGA proteins have been proposed to be involved in long-range enhancer-promoter interactions. For example, inducible expression of the chicken  $\beta$ -globin gene requires a functional interaction between a specialized TATA-box motif located at -30 bp in the gene's 5' promoter and a distal 3' enhancer element located  $\pm 1.9$  kb downstream of the start of the  $\beta$ -globin's coding region (35). Bagga *et al* demonstrated that HMGA1 proteins mediate chromatin looping and long-range interactions of these widely spread regulatory regions to regulate the chicken  $\beta$ -globin gene (36). The HMGA proteins are also implicated in the chromatin remodelling that occurs in the promoter region of the human *IL-2R $\alpha$*  gene (37). Transcriptional activation of the *IL-2R $\alpha$*  gene involves more than restructuring of its proximal promoter region, as discussed in the previous section. To effect maximal transcription of the *IL-2R $\alpha$*  gene following T-cell activation, distinct gene regulatory regions, which are positioned along a stretch of more than 12 kb of genomic DNA, have to be temporally and coordinately controlled. HMGA1 binding sites are known to be present in all of the known *IL-2R $\alpha$*  gene regulatory regions and HMGA1 proteins physically interact with many of the transcription factors involved.

**Enhanceosome formation.** It is believed that specific gene expression patterns are controlled by massive molecular aggregates termed transcription complexes. These complexes, which assemble over the regulatory region of a gene, are composed of four classes of proteins that together govern the specificity and activity of transcription. The first class of proteins is composed of activators. Activators bind co-operatively to promoters and enhancers to assemble into nucleoprotein complexes called enhanceosomes. The second class of proteins is composed of co-activators, which are recruited to the gene by the enhanceosome. The co-activators then interact with the third class of proteins: general factors such as RNA polymerase, which carry out the catalytic process of transcription initiation and elongation. Co-activators also interact with a fourth class of factors: elements that modify and remodel chromatin. These elements thus assist in enhanceosome

assembly by bending DNA to allow the cooperative interactions to occur. Seen the unusual structural characteristics of the HMGA proteins, they could easily fit in this fourth class of proteins.

HMGA proteins are indeed known to be able to orchestrate the regulation of gene expression by altering DNA conformation and recruiting transcription factors to DNA. One of the best-characterized mechanisms of gene regulation in which HMGA1 proteins are involved, is that of the virus inducible enhancer of the human  $\beta$ -interferon gene [*IFN- $\beta$* ; (38)]. The DNA-protein complex formed on the *IFN- $\beta$*  promoter serves as a stereo-specific platform involved in recruiting RNA polymerase II and other co-factors necessary for transcription initiation. In its basal configuration, *IFN- $\beta$*  enhancer DNA is in an unfavourable conformation to allow recognition by its cognate transcription factors. However, after viral infection, HMGA1 enhances the interaction of transcription factors (such as NF- $\kappa$ B, ATF-2/c-jun, and IRF-1/3/7) with their binding sites by adhering to AT-rich sequences and so reversing the curvature of the DNA. HMGA1 proteins also engage in protein-protein interactions with the bound transcription factors, thereby stabilizing the enhanceosome (also see above). In a next step, the transcriptional coactivator complex GCN5/PCAF and the RNA polymerase II are recruited to complete the complex assembly process of the massive molecular aggregate resulting in the initiation of transcription. A similar mechanism is seen in the HMGA1-dependent regulation of *NOS2* (39).

*Protein-protein interactions.* As discussed above, HMGA proteins are known to regulate transcription of many genes by binding to DNA and hereby altering the structure of the DNA, allowing cooperative recruitment of gene activators. Together with the HMGA proteins, these activators assemble into a higher order nucleoprotein complex, called the enhanceosome. Indeed, the assembly and function of the enhanceosome requires a complex network of protein-DNA and protein-protein interactions as is shown for the *IFN- $\beta$*  gene. These protein-protein interactions involve binding between HMGA proteins and other nuclear proteins, most of which are transcription factors (reviewed in refs. 4 and 40). Upon binding, HMGA proteins induce structural changes in their bound protein substrates. Different transcription factors interact with different parts of the HMGA proteins. These interaction sites are distributed along their entire length. Some of the transcription factors interact with large areas of the HMGA proteins [e.g. ATF-2/c-Jun (13)], while others interact with a much more restricted portion. In any case, the site of interaction includes part, or the whole, of one or more AT-hook motifs, plus flanking regions. The region between the second and third AT-hook of the HMGA1 protein, for example, has the most identified interacting partners. The NF- $\kappa$ B p50 subunit is known to interact with this region in both the HMGA1 and HMGA2 proteins (14). The transcriptional repressor p120<sup>E4F</sup> also interacts with HMGA2 via this region (41).

Via direct protein-protein interactions with transcription factors, HMGA proteins can also influence gene transcription without first binding to DNA. HMGA1, for example, is able to enhance binding of serum response factor (SRF) to DNA by

interacting with SRF, leading to SRF-dependent transcription (15). HMGA2, on the other hand, can for example interact with the E1A-regulated transcriptional repressor p120<sup>E4F</sup>. This interaction results in disruption of p120<sup>E4F</sup> binding to the cAMP response element (CRE) site of the promoter and subsequent activation of *cyclin A* gene transcription (41).

*Virus integration and expression.* From the above, it is clear that the HMGA proteins are implicated in a wide range of cellular processes including chromosome changes during the cell cycle, chromatin remodelling, regulation of cellular gene transcription, DNA replication, and DNA repair. In addition, the HMGA proteins also seem to be important in integration of retroviruses into host genomes, as well as viral gene transcription. The HMGA proteins have been shown to act as co-factors in the infectious process of HIV-1 (42), Moloney murine leukemia virus (MoMuLV) (43,44), and avian sarcoma virus (ASV) (45) by facilitating integration of double stranded linear cDNA copies of their genomes into host cell chromosomes (46,47). HMGA proteins appear to be associated with pre-integration complexes (PICs). It has been proposed by Li *et al* that binding of monomeric HMGA proteins to multiple AT-rich sites on retroviral cDNAs in PICs leads to DNA compaction and the formation of active integrase-cDNA complexes (43). The HMGA proteins are also employed as host cell-supplied factors involved in controlling transcriptional expression of a number of viral genes. They bind to AT-rich regions of viral promoter/enhancer regions and, in cooperation with both viral copies and host proteins, control viral gene transcription. Examples include regulation of the early and late genes of the human JC papovavirus (48); the *IE-3* gene (which codes for the immediate early protein ICP4) and the latency active promoter (LAP-2) of herpes simplex virus-1 (HSV-1) (49); and the human papillomavirus type 18 (HPV18) oncogenes (50,51). In these examples, HMGA1 is involved in the formation of the enhanceosome to activate transcription. HMGA1 can also cooperate with the chromatin remodelling complex (the hSWI/SNF complex), which is situated in the 5' long terminal repeat (LTR) promoter/enhancer of the human HIV-1 virus (52).

The role played by the HMGA proteins in retroviral integration and viral gene transcription mechanistically mimics the normal function(s) of these proteins in eukaryotic cells.

*Effect of post-translational modifications.* The HMGA proteins are among the most highly adducted proteins in the nucleus, exhibiting complex patterns of phosphorylations, acetylations, methylations and possibly other modifications [e.g. sumoylation (53), ribosylation (54)]. It is obvious that these secondary biochemical modifications may profoundly alter the affinity of HMGA proteins to bind to both DNA and proteins and consequently affect their biological activities. Most studies regarding post-translational modifications have focused on the HMGA1 protein, which will therefore be discussed in more detail (also see Fig. 3).

With regard to phosphorylation, HMGA proteins are amongst the most highly phosphorylated proteins in the nucleus. For example, as mentioned in the part about the role of HMGA proteins and chromatin structure, cdc2 kinase



methylation in this region thus probably plays an important role in HMGA functions. All findings on methylation of the HMGA1 proteins on different Arg residues were very recently confirmed and extended by Zou *et al* (67).

**Differential functions for HMGA1a, HMGA1b and HMGA2?** The HMGA family members, being HMGA1a/b/c and HMGA2, possess a very high sequence homology (Fig. 2). The HMGA1a and HMGA1b proteins are nearly identical in amino acid sequence, the only difference being an internal deletion of 11 amino acids, as stated above. Therefore, one might ask the question to which extent they control the expression of the same set of genes. It is clear that the HMGA proteins are indeed able to regulate transcription of some common target genes. For example, both HMGA1 and HMGA2 proteins regulate the promoter activity of the *IFN- $\beta$*  and *ERCC1* genes. However, it appears that there are some differences in binding properties and affinities of the HMGA1 and HMGA2 proteins to the promoters of these genes. For example, HMGA2 appears to have a higher affinity for the *ERCC1* gene in comparison to HMGA1 (32). Other examples include IL-2 and IL-2R $\beta$  where HMGA1 has a higher affinity than HMGA2 (68) and IL-15 and IL-15R where HMGA2 has a higher affinity than HMGA1 (69). However, HMGA1 and HMGA2 also show differences in target gene selection. The *BRCA1* gene, for example, is negatively regulated by HMGA1 and not by HMGA2 (70).

It was long believed that the HMGA1a and HMGA1b proteins would be biologically interchangeable, since the only difference between them is an internal deletion of 11 amino acids. However, more and more reports now demonstrate that there are in fact some differences between the two isoforms. Cmarik *et al* undertook a study in which they compared the levels of the two proteins in response to treatment of cells with the tumor promoting agent 12-O-tetradecanoylphorbol acetate (TPA). With their experiments, they were able to show that differential induction of HMGA1b, relative to HMGA1a, was important for the TPA-induced neoplastic transformation process (71). Direct confirmation that HMGA1a and -A1b could indeed have different functions *in vivo*, also came from studies performed by Reeves *et al*. These experiments showed that tetracycline-regulated induction of HMGA1b in the non-metastatic MCF-7 mammary epithelial cells caused them to progress much more rapidly to a metastatic and highly malignant phenotype than did induced over-expression of HMGA1a. Furthermore, microarray studies comparing the gene expression profile in HMGA1a to that in HMGA1b overexpressing MCF-7 cells, clearly illustrated that the two isoforms differentially regulate specific genes in the transgenic MCF-7 cells (72).

Since the amino acid composition of the DNA binding domains of the HMGA1 and HMGA2 proteins is nearly identical, the difference in target gene selection and/or binding affinities must most likely be explained by other regions of these proteins. It appears that one reason for their ability to exert dissimilar functions is due to the different spacing of the AT-hooks along the protein molecules (Fig. 4). This allows the different HMGA proteins to establish interactions with alternatively spaced AT-rich DNA stretches. From studies concerning possible differential *in vivo* modifications,

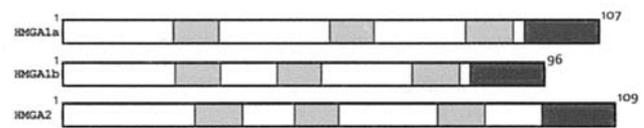


Figure 4. Schematic overview of HMGA proteins emphasizing the differently spaced AT-hooks. AT-hook (light grey boxes), C-terminal tails (dark grey boxes) (reproduced with permission from ref. 169).

it is also clear that HMGA1a and HMGA1b proteins might participate in different cellular processes based on their unique collections of *in vivo* post-translational modifications (56,64). In the next section it will become clear that these differences in target gene selection and/or binding affinity by the HMGA family members result in different cellular functions.

## 5. Cellular functions of the HMGA proteins

**HMGA proteins in growth and differentiation.** In general, the development of an organism consists of a few fundamental aspects. One of them is 'growth'. Classical genetic studies in mice have isolated four viable, spontaneous mutants disrupted in growth, leading to dwarfism (73). Three of the four phenotypes could be explained by aberrations in the growth hormone insulin-like growth factor endocrine pathway. One of the phenotypes, called 'pygmy', appeared to be the result of a major disruption at a particular spot in the genome. The gene affected was identified as being the *hmga2* gene (74). The same scientists then generated a mouse null mutant of *hmga2* to exclude the possibility that another gene might be responsible for the pygmy phenotype. Homozygous *hmga2*<sup>-/-</sup> mice revealed the classical features of the pygmy phenotype including reduced birth weight, craniofacial defects (shortened head), and an adult body weight of approximately 40% of that of wild-type littermates (74,75). Further insight into the function of *hmga2* was obtained by studying the growth characteristics of fibroblasts derived from *hmga2* null mice. These fibroblasts showed a 4-fold less proliferation capacity compared to fibroblasts derived from wild-type mice (74,75). In 12.5-dpc embryos of both *hmga2*-deficient mice strains (pygmy and *hmga2*<sup>-/-</sup> mice), it was found that the *insulin-like growth factor II mRNA binding protein 2 (Imp2)* gene was significantly down-regulated as compared to wild-type embryos (76). *Imp2* is an mRNA-binding protein which is involved in temporal and spatial control of gene expression at the level of mRNA rather than at the level of gene transcription. In this regard, it is tempting to speculate that the effect of *hmga2* on cell proliferation during embryonic development involves *imp2*. More recently it was shown by our group that *Imp2* is regulated by HMGA2 via an intronic regulatory element, in cooperation with NF- $\kappa$ B (77).

These data, together with the observation that HMGA proteins are rapidly induced in quiescent normal cells following exposure to growth stimulatory factors and thus classifying the HMGA proteins as 'delayed early response genes' (78), clearly suggest involvement of the HMGA genes in growth regulation.

*In situ* hybridization analysis of both *hmga1* and *hmga2* showed a very similar time course of expression during fetal

life. High levels of expression are observed in nearly all tissues in the earliest stages analyzed: 8.5 dpc for *hmgal* (79) and 9.5 dpc for *hmg2* (80). During later stages of development, the expression patterns become more restricted to specific organs and tissues. For *hmgal*, the expression is confined to tissues of ectodermal, mesodermal and endodermal origin. *Hmg2* expression on the other hand, is mainly confined to mesodermal-derived tissues. At stages later than 15.5 dpc, Northern blot analysis showed low expression of *hmgal*, and no detectable expression of *hmg2* (74,81). *Hmgal* expression is observed at low levels in adult tissues of both human and mice (79), while *hmg2* expression is very low if detectable at all (82).

Another aspect of fundamental importance for embryonic development is the establishment of the correct nature of cell types - or 'differentiation'. It appears that while *hmg2* is mainly considered to be engaged in growth, *hmgal* is more than a proliferation factor alone. For example, *hmgal* is expressed in the cortical plate of the telencephalon and the spinal cord at 17.5 dpc. These tissues are composed mainly of differentiated, non-dividing cells. *Hmgal* also shows a dramatically decreased expression during mouse teratocarcinoma differentiation (83) and an increased expression as cells dedifferentiate due to the expression of *v-mos* or *v-Ras* oncogenes (84). Furthermore, *hmgal* has been shown to exert a negative effect on the proliferation of adipocytic 3T3-L1 cells and a positive effect on differentiation of these cells (85). The mouse pre-adipocyte 3T3-L1 cell line is a well characterized differentiation model. After hormonal stimulation, the 3T3-L1 cells can differentiate into adipocytes (86). The *hmgal* proteins seem to exert this role in adipocytic differentiation, in part, by physically interacting with C/EBP proteins, hereby functionally cooperating in their transcriptional activity. The function of C/EBP proteins is required to trigger the expression of adipocyte-specific genes. Conversely, *hmg2* has a role in the proliferation step which precedes - and also is required for - the later phases of adipocyte differentiation (87).

In an attempt to clarify the role of *hmgal* proteins in embryonic development, Battista *et al* used ES cells carrying disruption of one or both alleles of the *hmgal* gene. It appeared that *hmgal* proteins are involved in different hematopoietic lineage commitment checkpoints. In *hmgal*<sup>-/-</sup> embryoid bodies (EBs), myeloid differentiation was shown to be impaired, while megakaryocytic and erythroid differentiation was shown to be increased (88). *Hmgal* proteins were also shown to play a role in determining the fate of the lymphoid cell common precursor: the number of T-cell precursors was greatly reduced in *hmgal*<sup>-/-</sup> cells, whereas the number of B-cell precursors was increased. Later, they showed that *hmgal* down-regulates *recombination activating gene 2* (*RAG2*), which is a key regulator of lymphoid differentiation, by directly inhibiting its promoter activity. They hereby suggest an additional mechanism for the modulation of lymphopoiesis by *hmgal* proteins (89).

Another group showed that both *hmgal* alleles are required for normal sperm development in mice (90). *Hmgal* proteins are also shown to have a critical role in heart development and growth, since *hmgal*-null mice show cardiac hypertrophy due to the direct role of *hmgal* on cardiomyocytic cell growth regulation (91).

Recently, it was proposed that the *hmg2* protein could also be involved in differentiation, namely in skeletal muscle differentiation of embryonic stem cells (92).

Together, these data strongly support a role for the HMGA proteins in the regulation of growth and differentiation, two fundamental aspects of development. Furthermore, it is clear that the HMGA proteins exert different biological roles. The *hmg2*<sup>-/-</sup> mouse for example shows a pygmy-phenotype which is not compensated for by the *hmgal* proteins. In fact, *hmgal* expression was not altered in the pygmy-mouse in comparison to wild-type mice (74). In addition, the *hmgal*<sup>-/-</sup> mouse shows a different phenotype than that of the *hmg2*<sup>-/-</sup> mouse: the former ones develop cardiac hypertrophy, myelo-lymphoproliferative disorders and diabetes (91,93).

#### *HMGA proteins in tumors*

*Overexpression of the HMGA proteins in tumors.* As mentioned above, *HMGA1* and *HMGA2* are hardly detectable in adult human tissues. The same holds true for *hmgal* and *hmg2* in mouse tissues. There are, however, some exceptions to the general repression of these genes in adult tissues. For example, there is a burst of synthesis of the HMGA1a protein during the immune response [(94,95); also see above] and in photoreceptor cells a constitutive expression of the HMGA1 proteins is observed (96). The most studied and remarkable observation, however, is that there is a generalized derepression of HMGA1 and often HMGA2 expression in neoplastic cells.

Overexpression of members of the HMGA family, in particular of HMGA1, was first observed in HeLa cells (7) and in cultured rat thyroid cells transformed by viral oncogenes (84). HMGA1 expression was also shown to correlate with the progressive transformation of rat prostate cells (97) and mouse mammary epithelial cells (98). In humans, overexpression of HMGA1 has been described in a large number of different types of cancer, all being malignant epithelial tumors. These include those of the following organs: prostate (99,100), colorectum (101-103), lung (104), breast (105,106), pancreas (107,108), and thyroid (109). In all these tumors, increasing concentrations of HMGA1 proteins are correlated with increasing degrees of malignancy or metastatic potential.

HMGA2 overexpression is also observed in various malignant tumor types: breast cancer (110), sarcomas (111), pancreas (112), oral squamous cell carcinomas (113), and non-small cell lung cancer (114). Furthermore, in well-differentiated liposarcomas (WD-LPS), it was shown that there exists a positive and statistically significant correlation between *HMGA2* and *IMP2* gene expression (77). As stated above, *Imp2* was identified as an HMGA2 target gene in 12.5 dpc embryos of two *hmg2*-deficient mice strains. In contrast, Brants *et al* showed that HMGA2Tr, which is frequently expressed in benign solid tumors (also see below), exerts a negative effect on *Imp2* gene expression (76). These observations point towards the potential significance of a deregulated HMGA2-IMP2 pathway in this kind of tumors.

The role of HMGA1 and HMGA2 overexpression in tumor development is further supported by different observations. Overexpression of HMGA1 and HMGA2 proteins led to transformation of Rat1a fibroblasts and human lymphoid CB33 cells (115,116). Quite recently, *hmg1a*, *hmg1b* and *hmg2* transgenic mice were developed. Both *hmg1b* and

*hmg2* transgenic mice develop mixed growth hormone/prolactin cell pituitary adenomas (69,117). However, the *HMGA1* gene has not yet been implicated in the generation of human pituitary adenomas. On the contrary, the *HMGA2* gene is found to be amplified in human prolactinomas, and this is associated with HMGA2 overexpression (118). A crucial step in the onset of pituitary adenomas in *hmg2* transgenic mice, and probably also in humans, is E2F1 activation (119,120). The transcriptional activity of the E2F family of transcription factors is crucial for the expression of several genes required to enter the S-phase of the cell cycle (121). Hmg2 acts by displacing HDAC1 from E2F1 target promoters, hereby recruiting histone acetylases which first acetylate the histones - relieving transcriptional repression - and then also acetylate E2F1 causing its stabilization (119,120). This model provides a very nice example of the mode of action of HMGA proteins to activate transcription and of the effect of post-translational modifications on their function (also see above). *Hmg1a* transgenic mice develop aggressive, highly penetrant lymphoid malignancy and *HMGA1a* is also overexpressed in human lymphoid leukemia (122,123). In addition to lymphoid malignancy, *Hmg1a* transgenic mice develop uterine tumors, which resemble human uterine adenocarcinomas (124). In most high-grade, but not in benign or most low-grade human uterine adenocarcinomas, significant overexpression of *HMGA1a* is detected (124). *HMGA1a* could thus serve as a useful marker for such aggressive uterine tumors.

*Rearrangements of HMGA1 and HMGA2 genes in human benign tumors.* Tumor type-specific chromosomal rearrangements resulting in the formation of a specific fusion protein have been shown to be causally related to the development of many tumor types. Rearrangements in the *HMGA* genes are found in various benign mesenchymal tumors such as lipomas (125,126), pleiomorphic adenomas of the salivary glands (127), uterine leiomyomas (128-130), angiomyxomas (131), pulmonary chondroid hamartomas (130,132), and endometrial polyps (133-135). Since most translocations involve the *HMGA2* gene, the focus of this section will mainly be on HMGA2.

For *HMGA2*, these translocations involve chromosome region 12q13-15. The chromosome breakpoints often occur in the large third intron of *HMGA2*, which separates the DNA-binding domains from the acidic domain (see above). The translocation thus results in a disrupted *HMGA2* gene wherein the AT-hooks of *HMGA2* become separated from the acidic tail and become fused to a variety of ectopic sequences. In lipomas, the preferred translocation partner was shown to be located in chromosome region 3q27-28. In this region, the LIM-domain containing lipoma preferred partner (*LPP*) was identified (136). In uterine leiomyomas, the recombinational repair gene *RAD51B* on chromosome 14q23-24 was identified as the preferential translocation partner of *HMGA2* (137). Some of the fusions involving *HMGA2* occur in frame. In that case, the fusion partner could have a specific role in tumor formation, for instance in the case of the *HMGA2-LPP* fusion. It is noteworthy that the *HMGA2-LPP* fusion protein is not specific for lipomas, but also has been detected in pulmonary chondroid hamartomas (138).

On the other hand, it is possible that the in frame fusion partner contributes only a few amino acids to the chimeric protein. The mitochondrial dehydrogenase gene (*ALDH2*), identified as a fusion partner of *HMGA2* in a primary uterine leiomyoma, for example, contributes less than 10 amino acids to the fusion protein (139). However, most of the fusions occur out of frame, and mostly a stop codon is encountered quickly so that only a few amino acids are added to the AT-hooks of *HMGA2*. An example is the *lumican* encoding gene (*LUM*), which is observed as a translocation partner of *HMGA2* in an osteosarcoma-derived cell line (140). All this suggests that the minimal requirement for tumorigenesis would be *HMGA2* activation due to rearrangements that leave intact at least exons 1, 2 and 3, which encode for the AT-hook domains. In conclusion, translocations involving 12q13-15 often result in the formation of a truncated form of *HMGA2*, called *HMGA2Tr*, mainly consisting of the three DNA-binding domains.

Rearrangements in other regions than the large third intron have also been described. For example, in uterine leiomyomas, breakpoints upstream and downstream of the *HMGA2* gene have been identified (141,142). The existence of breakpoints outside the gene could mean that transcriptional activation, as a consequence of its juxtaposition to enhancer elements in the translocation partner, is sufficient to induce tumor development.

Recent, new insights indicate that the removal of the 3' untranslated region (UTR) by chromosomal rearrangements may contribute to tumorigenesis. The 3' UTR contains multiple putative *let-7* binding sites. *Let-7* was one of the first identified micro-RNA (miRNA) genes, the expression of which seems to be developmentally regulated in both human (143) and mice (144). Intriguingly, the expression of *let-7* during development is inversely correlated with that of *HMGA2*, leading to the hypothesis that *let-7* might act as a repressor of *HMGA2*. In addition, *let-7* seems to be deregulated in uterine leiomyomas, in which *HMGA2* translocations - resulting in the removal of the 3' UTR - are often observed. This type of post-transcriptional regulation by miRNAs, as well as its functional importance is described below in further detail.

The role of *HMGA2* disruption as the initiating event in tumor development is, however, somewhat controversial (145). Still, there is evidence supporting a causal role for the *HMGA2* protein in the onset of benign neoplasia. *In vitro* experiments show that both *HMGA2Tr* and *HMGA2/LPP* are able to induce neoplastic transformation of NIH-3T3 fibroblasts (146). In addition, Crombez *et al* showed that in a number of lipomas, *HMGA2/LPP* and *HMGA2* are co-expressed, and that *HMGA2* augments the transactivation functions of *HMGA2/LPP* (147). These results support the concept that the transactivation functions of the *HMGA2/LPP* transcription factor contribute to lipomagenesis. *In vivo* evidence includes observations in both mouse and human. *Hmg2Tr* transgenic mice developed by the group of Fusco present gigantism, associated with lipomatosis (148). They also develop natural killer NK-T/NK cell lymphomas, which are very similar to the NK-T/NK cell lymphomas seen in *hmg1b* transgenic mice. The induction of NK-T/NK cell lymphomas in these transgenic mice is probably due to overexpression of *IL-2* and *IL-15*, which are shown to be target

genes of HMGA2, HMGA2Tr and HMGA1 (68,69). *Hmga2Tr* transgenic mice made independently by Arlotto *et al* develop adiposity and display an abnormal high incidence of lipomas (149). In 2005, an 8-year-old male was described with a *de novo* pericentric chromosomal inversion of chromosome 12, resulting in the expression of a truncated form of HMGA2 only containing the DNA-binding AT-hooks. The patient showed somatic overgrowth, multiple lipomas, advanced endochondral bone and dental aging and a cerebellar tumor (150). Together, these findings seem to suggest that the DNA-binding domains of HMGA2 are sufficient to predispose to various benign neoplasms.

HMGA2 rearrangements have not only been identified in benign neoplasms, but also in several types of malignant mesenchymal tumors such as well-differentiated liposarcomas (151), and inflammatory myofibroblastic tumors (152).

For HMGA1, rearrangements involve chromosomal region 6p21.3. Breakpoints in most tumors seem to be located 3' of HMGA1 (130), although breakpoints located 5' (130) or intragenic (153,154) have also been reported.

**HMGA proteins in other critical diseases.** The proper control of constitutive and inducible gene expression is critical for the physiologic homeostasis of a complex organism, like for example, the human body. Abnormal regulation of gene expression is therefore likely to contribute to a pathophysiological process for many diseases. The manner in which genes are regulated may for instance determine the aggressiveness of a disease or the ability of a patient to survive a critical illness. This not only holds true for cancer, as described above, but also for other critical illnesses. Architectural transcription factors, such as the HMGA proteins, thus constitute interesting targets in disease in which such genes are expressed inappropriately.

As a first example, HMGA proteins are known to contribute to the regulation of many genes that are involved in the inflammatory response. HMGA1 regulates genes which contribute to the proliferation, differentiation, antigen presentation, and recruitment of inflammatory cells, as well as to cytokine and inflammatory mediator production that participate to the inflammatory response. These genes include: *nitric oxide synthase 2 (NOS2)*, *cyclo-oxygenase 2 (Cox-2)*, *E-selectin*, *Immunoglobulin E (IgE)*, *Interleukin 4 (IL-4)*, *IL-2*, *granulocyte macrophage-colony stimulating factor (GM-CSF)*, and *IFN- $\beta$*  (reviewed in ref. 39). When the host's immune cells - including macrophages - are activated, they release pro-inflammatory cytokines, such as IL-1 $\beta$  and tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ). Uncontrolled release of these pro-inflammatory cytokines would lead to hypotension and shock. Indeed, an overpowering inflammatory response, owing to severe infection, results in sepsis. Since HMGA1 is known to contribute to the regulation of a number of genes involved in the inflammatory response, and since many of the signals responsible for gene expression involve an inflammatory response, it is believed that HMGA1 proteins could have a potential role in the pathophysiology of sepsis (39). Based on several pieces of information, it has also been hypothesized that impairment of HMGA1 function might be responsible for some cases of human immunodeficiency: i) *hmgal*-null mice develop B-cell lymphomas (91), ii) *hmgal* overexpressing

mice develop T-cell lymphomas (69), iii) *hmgal* has been shown to regulate RAG2 expression in a negative fashion (89), iv) inactivating mutations in *RAG2* are responsible for the 'Omen syndrome', in which no circulating mature B-cells are found, while a large number of poorly functional T-lymphocytes could be detected (155), and v) overexpression of RAG proteins may contribute to some cases of human immunodeficiency (156) and lymphocytic leukemia.

Another example of a critical illness with an important role for HMGA1 is type 2 diabetes mellitus. Type 2 diabetes mellitus is a widespread disease, affecting millions of people worldwide. The cause of this metabolic disorder is largely unknown. A few years ago, four type 2 diabetes patients were reported with a markedly reduced expression of HMGA1. This appeared to negatively affect insulin receptor expression and function in insulin target cells and tissues. This defect in HMGA1 expression may have induced the insulin resistance and type 2 diabetes in these patients (93). It was already shown before that HMGA1 is required for proper *INSR* gene transcription (157,158). The same group also developed *hmgal* knock-out mice. These mice showed a considerably decreased insulin receptor expression in the major targets of insulin action, largely impaired insulin signalling, and severely reduced insulin secretion, causing a phenotype characteristic of human type 2 diabetes (93).

Besides a diabetes phenotype, *hmgal*<sup>+/-</sup> and *hmgal*<sup>-/-</sup> mice also develop cardiac hypertrophy. This is due to the role of *hmgal* on cardiomyocytic cell growth regulation, via increased levels of class II calcium/calmodulin-dependent protein kinase (CaMKII) - which is normally repressed by *hmgal* (91).

## 6. Regulation of HMGA expression

*HMGA* gene expression can be regulated in response to a large array of cellular stimuli. Many growth factors are able to induce *HMGA* gene expression through mechanisms common to the delayed early response of gene expression (reviewed in ref. 159). They are thus induced after a delay of a few hours after growth factor addition and require new protein synthesis (78). *HMGA1* has also been shown to be induced by the tumor-promoting agent 12-O-tetradecanoyl phorbol-13-acetate (TPA) (160).

Moreover, *HMGA* expression can be repressed by activation of a differentiation program. For example, in neuroblastoma and in embryonic carcinoma cells, *HMGA1/2* expression can be repressed by retinoic acid (87,161,162). In addition, IL-1 $\beta$ , endotoxin and IFN- $\beta$  - all mediators of the inflammation program - can induce *HMGA1* expression. As mentioned above, the condition most consistently associated with deregulated *HMGA1/2* expression, is cellular transformation. Nevertheless, the question of how *HMGA1* and *HMGA2* expression is deregulated in tumors is not clear. In what follows, possible mechanisms of *HMGA1/2* deregulation in cancer will be discussed.

**Transcriptional (de)regulation of HMGA1.** Transcriptional deregulation is probably a major mechanism involved in the aberrant expression of *HMGA1* in tumors. However, little is known about transcriptional regulation of *HMGA1* in both

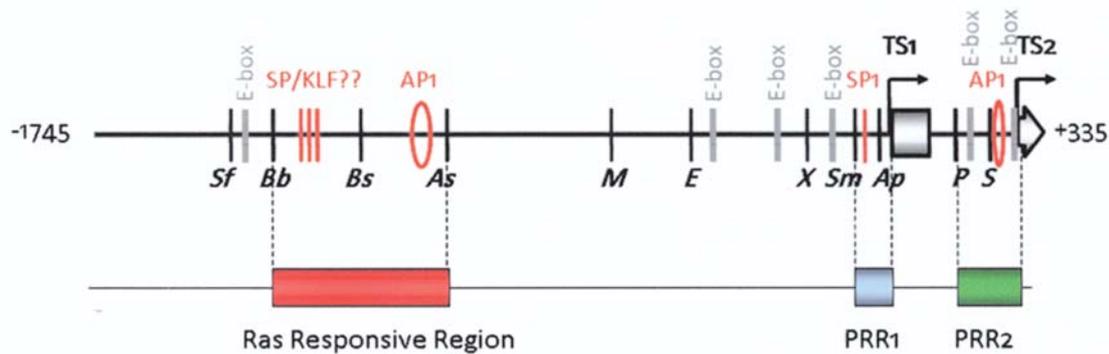


Figure 5. Schematic overview of the *HMGA1* promoter. PRR1, proximal regulatory region 1; PRR2, proximal regulatory region 2; TS1, transcriptional start site 1; TS2, transcriptional start site 2; restriction enzyme recognition sites used to make 5'-deletion constructs of the *HMGA1* promoter (italics): Sf (*SfiI*), Bb (*BbrPI*), Bs (*Bsu36I*), As (*AscI*), M (*MluNI*), E (*EcoRI*), X (*XhoI*), Sm (*SmaI*), A (*ApaI*), P (*PacI*) and S (*SspI*). A detailed description of how the different regions were defined, can be found in ref. 164. The transcription factor binding sites, for which an involvement in *HMGA1* regulation is shown, are indicated.

normal cells as well as in tumor cells. A partial explanation for this comes from the extremely complex structure and transcriptional regulation of both the mouse and human *HMGA1* gene. It comprises four different transcription start sites, of which the first two are considered to be the major ones (160).

Only a few *cis*-acting elements have been characterized on the *HMGA1* promoter. A schematic overview of the human *HMGA1* promoter and the transcription factor binding sites for which an involvement in *HMGA1* regulation is shown, is depicted in Fig. 5. An *API*-binding site located immediately before start site 2 of the human *HMGA1* promoter was shown to be required for TPA induction in K562 cells (160). Furthermore, *hmgal* is a direct target of *API* in Rat1a cells via binding of *c-Jun* to an *API*-binding site located at -1337 bp of the mouse *hmgal* promoter (163). An *E-box* located at -1353 and -1337 bp of the human and mouse *HMGA1* promoter respectively, is essential for *c-myc*-dependent induction of *HMGA1* expression in Burkitt's lymphoma cell lines and HEK293 cells (116). In addition, *HMGA1* expression is also associated with *MYCN* amplification in primary human neuroblastomas. *HMGA1* appears to be a direct *MYCN* transcriptional target, wherein *MYCN* mediates *HMGA1* transactivation by multiple *cis*-acting elements located near the first three transcription start sites (162).

The work done by Wood *et al*, indicates that activation of *HMGA1* expression by the *c-myc* oncogene is essential for the appearance of the fully transformed Burkitt's lymphoma phenotype. Indeed, inactivation of *HMGA1* expression reversed the transformed phenotype (116). Still, it is clear that *c-myc*-induced *HMGA1* expression cannot represent a ubiquitous means of reactivating its expression. Indeed, we reported recently that the *API*-binding site before start site 2 (in the so-called Proximal Regulatory Region 2), together with an *SP1*-binding site before start site 1 (in Proximal Regulatory Region 1) are important for basal transcription of the *HMGA1* gene (164). In addition, we have shown that the *HMGA1* promoter is strongly inducible by oncogenic *Ras*, via a distal regulatory region (Fig. 5). An *API*-binding site and three *SP1*-like-binding sites are responsible for this inducible activity. The *E-box* located at -1353 bp was not of importance for the oncogenic *Ras*-induced expression. An even more

convincing finding for a role of oncogenic *Ras* in the regulation of *HMGA1* in cancers, was the discovery that *HMGA1* upregulation in the HCT116 colon cancer cell line is abolished when the mutated *Ras* allele is removed from these cells. These data clearly show that it is thus more than likely that not only *c-myc*, but also other oncogenic pathways directly target *HMGA1* expression in neoplasia.

Further detailed analysis of the *HMGA1* promoter could help to identify other *cis*-acting elements and *trans*-acting factors responsible for *HMGA1* deregulation in so many different types of cancer.

**Regulation of *HMGA2* expression.** Experiments performed by Ayoubi *et al* demonstrated that the *HMGA2* promoter also contains multiple transcription initiation sites, which are differentially utilized. In addition, they show that *HMGA2* 5'-flanking sequences had constitutive promoter activity in cell lines which express high levels of *HMGA2*, and in cell lines which do not express *HMGA2* (87). It is thus possible that *HMGA2* is regulated by negative regulatory elements, which were not included in the described experiments, and/or that *HMGA2* could be regulated at the post-transcriptional level.

Concerning the latter, very recent, new insights have emerged. It was always presumed that the loss of its C-terminal region, resulting from for example chromosomal translocations concerning the *HMGA2* gene, would be the cause of oncogenic transformation. However, chromosomal rearrangements in some tumors leave the ORF intact, but instead disrupt the 3' UTR. This is associated with overexpression of the wild-type *HMGA2* protein (125,127). In addition, luciferase reporter experiments in which large portions of the *hmgal2* 3' UTR are used, show repression of luciferase activity (165). Moreover, transgenic mice overexpressing wild-type *hmgal2* have similar phenotypes as those expressing *hmgal2Tr*: both develop abdominal lipomatosis, lymphomas, pituitary adenomas, and lung adenomas (68,117,148). It is, therefore, now believed that transformation by *HMGA2* might be caused by the loss of repressive elements in its 3' UTR. Within the last few years, miRNAs have emerged as important regulators of both transcriptional and post-transcriptional gene silencing. miRNAs are RNAs of approximately 22 nucleotides in length that typically bind to the 3' UTRs of specific mRNA

targets and repress their expression at the post-transcriptional level. Aberrant miRNA expression can contribute to tumorigenesis. Recently, two independent groups showed that the *HMGA2* transcript can be regulated by *let-7* miRNA. Disruption of the *let-7* repression of *HMGA2*, by inhibiting *let-7* using a 2'-O-methyl antisense oligonucleotide against *let-7*, enhances cell proliferation in H1299 (166) cells and causes transformation in NIH-3T3 cells (167). *Let-7* miRNA thus might act as a repressor of *HMGA2*. In conclusion, losing miRNA-directed repression of an oncogene could provide a mechanism for tumorigenesis. This mechanism could also apply for *HMGA1*, since numerous *HMGA1* transcripts with a truncated 3' UTR have been reported in tumors (130,168).

## 7. Conclusion

Together, these observations establish an extremely important role for the HMGA proteins in cancer and other diseases. It is, therefore, of crucial importance to have insight in the mechanism(s) behind the elevated regulation of expression of *HMGA1*, but also in the exact role played by the HMGA proteins in the development of a variety of tumors and other diseases.

In an era where diseases such as cancer, diabetes, and obesity are prevailing, proteins like those of the HMGA family and the complex context in which they operate, will certainly be of interest to study in the context of pharmacogenomics. Pharmacogenomics holds the promise that drugs might one day be custom-made for individuals and adapted to each person's own genetic make-up. The anticipated benefit of pharmacogenomics is that pharmaceutical companies will be able to create drugs interfering specifically with the proteins, enzymes, or RNA molecules associated with genes which are linked to diseases. This will profoundly affect drug discovery strategies and enable drug makers to develop therapeutic strategies more specifically targeted to a particular disease of an individual or a group of individuals.

## Acknowledgements

This work was supported by the 'Geconcerteerde Onderzoekacties' 2002/010, the L317105N grant of the 'Effel Stichting' of the 'Fonds voor Wetenschappelijk Onderzoek Vlaanderen' (FWO), the Cancer Research Program of Fortis Bank Insurance 2002-2005, the Foundation for Biochemical and Pharmaceutical Research and Education, and the Belgian Federation against Cancer: Foundation of Public Interest. Isabelle Cleynen is an Aspirant fellow of the 'Fonds voor Wetenschappelijk Onderzoek Vlaanderen' (FWO). Many thanks to Dr J. Brants (current address: Human Pathology, Institut de Génétique et Biologie Moléculaire et Cellulaire (IGBMC), CNRS/INSERM/ULP, Illkirch Cedex, France) for letting us use and adapt Figs. 1, 2, 3 and 4.

## References

1. Alberts BDB, Hopkin K, Johnson A, Lewis J, Raff M, Roberts K and Walter P: Chapter 5: DNA and chromosomes: 169-194. In: Essential Cell Biology. 2nd edition. Garland Science Textbooks, p896, 2003.
2. Bustin M and Reeves R: High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Prog Nucleic Acid Res Mol Biol* 54: 35-100, 1996.
3. Bustin M: Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol* 19: 5237-5246, 1999.
4. Reeves R and Beckerbauer L: HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta* 1519: 13-29, 2001.
5. Bustin M: Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem Sci* 26: 152-153, 2001.
6. Goodwin GH, Sanders C and Johns EW: A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. *Eur J Biochem* 38: 14-19, 1973.
7. Lund T, Holtlund J, Fredriksen M and Laland SG: On the presence of two new high mobility group-like proteins in HeLa S3 cells. *FEBS Lett* 152: 163-167, 1983.
8. Friedmann M, Holth LT, Zoghbi HY and Reeves R: Organization, inducible-expression and chromosome localization of the human HMG-I(Y) non-histone protein gene. *Nucleic Acids Res* 21: 4259-4267, 1993.
9. Pedulla ML, Treff NR, Resar LM and Reeves R: Sequence and analysis of the murine Hmg1y (*Hmg1*) gene locus. *Gene* 271: 51-58, 2001.
10. Johnson KR, Cook SA, Ward-Bailey P, Bustin M and Davisson MT: Identification and genetic mapping of the murine gene and 20 related sequences encoding chromosomal protein HMG-17. *Mamm Genome* 4: 83-89, 1993.
11. Nagpal S, Ghosh C, DiSepio D, Molina Y, Sutter M, Klein ES and Chandraratna RA: Retinoid-dependent recruitment of a histone H1 displacement activity by retinoic acid receptor. *J Biol Chem* 274: 22563-22568, 1999.
12. Frank O, Schwanbeck R and Wisniewski JR: Protein footprinting reveals specific binding modes of a high mobility group protein I to DNAs of different conformation. *J Biol Chem* 273: 20015-20020, 1998.
13. Yie J, Liang S, Merika M and Thanos D: Intra- and intermolecular cooperative binding of high-mobility-group protein (IY) to the beta-interferon promoter. *Mol Cell Biol* 17: 3649-3662, 1997.
14. Noro B, Licheri B, Sgarra R, *et al*: Molecular dissection of the architectural transcription factor HMGA2. *Biochemistry* 42: 4569-4577, 2003.
15. Chin MT, Pellacani A, Wang H, Lin SS, Jain MK, Perrella MA and Lee ME: Enhancement of serum-response factor-dependent transcription and DNA binding by the architectural transcription factor HMG-I(Y). *J Biol Chem* 273: 9755-9760, 1998.
16. Reeves R and Wolffe AP: Substrate structure influences binding of the non-histone protein HMG-I(Y) to free nucleosomal DNA. *Biochemistry* 35: 5063-5074, 1996.
17. Nissen MS and Reeves R: Changes in superhelicity are introduced into closed circular DNA by binding of high mobility group protein I/Y. *J Biol Chem* 270: 4355-4360, 1995.
18. Hill DA, Pedulla ML and Reeves R: Directional binding of HMG-I(Y) on four-way junction DNA and the molecular basis for competitive binding with HMG-I and histone H1. *Nucleic Acids Res* 27: 2135-2144, 1999.
19. Wolffe AP: Architectural transcription factors. *Science* 264: 1100-1101, 1994.
20. Laemmli UK, Kas E, Poljak L and Adachi Y: Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. *Curr Opin Genet Dev* 2: 275-285, 1992.
21. Saitoh Y and Laemmli UK: Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell* 76: 609-622, 1994.
22. Saitoh Y and Laemmli UK: From the chromosomal loops and the scaffold to the classic bands of metaphase chromosomes. *Cold Spring Harb Symp Quant Biol* 58: 755-765, 1993.
23. Disney JE, Johnson KR, Magnuson NS, Sylvester SR and Reeves R: High-mobility group protein HMG-I localizes to G/Q- and C-bands of human and mouse chromosomes. *J Cell Biol* 109: 1975-1982, 1989.
24. Zhao K, Kas E, Gonzalez E and 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* 12: 3237-3247, 1993.
25. Kas E, Poljak L, Adachi Y and Laemmli UK: A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin. *EMBO J* 12: 115-126, 1993.
26. Radic MZ, Saghbini M, Elton TS, Reeves R and Hamkalo BA: Hoechst 33258, distamycin A, and high mobility group protein I (HMG-I) compete for binding to mouse satellite DNA. *Chromosoma* 101: 602-608, 1992.

27. Nissen MS, Langan TA and Reeves R: Phosphorylation by cdc2 kinase modulates DNA binding activity of high mobility group I non-histone chromatin protein. *J Biol Chem* 266: 19945-19952, 1991.
28. Strauss F and Varshavsky A: A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* 37: 889-901, 1984.
29. De Griijl FR, van Kranen HJ and Mullenders LH: UV-induced DNA damage, repair, mutations and oncogenic pathways in skin cancer. *J Photochem Photobiol B* 63: 19-27, 2001.
30. Adair JE, Kwon Y, Dement GA, Smerdon MJ and Reeves R: Inhibition of nucleotide excision repair by high mobility group protein HMGA1. *J Biol Chem* 280: 32184-32192, 2005.
31. Maloney SC, Adair JE, Smerdon MJ and Reeves R: Gene-specific nucleotide excision repair is impaired in human cells expressing elevated levels of high mobility group A1 non-histone proteins. *DNA Repair (Amst)* 6: 1371-1379, 2007.
32. Borrmann L, Schwanbeck R, Heyduk T, Seebeck B, Rogalla P, Bullerdiek J and Wisniewski JR: High mobility group A2 protein and its derivatives bind a specific region of the promoter of DNA repair gene ERCC1 and modulate its activity. *Nucleic Acids Res* 31: 6841-6851, 2003.
33. Adair JE, Maloney SC, Dement GA, Wertzler KJ, Smerdon MJ and Reeves R: High-mobility group A1 proteins inhibit expression of nucleotide excision repair factor xeroderma pigmentosum group A. *Cancer Res* 67: 6044-6052, 2007.
34. Reeves R, Leonard WJ and Nissen MS: Binding of HMG-I(Y) imparts architectural specificity to a positioned nucleosome on the promoter of the human interleukin-2 receptor alpha gene. *Mol Cell Biol* 20: 4666-4679, 2000.
35. Fong TC and Emerson BM: The erythroid-specific protein cGATA-1 mediates distal enhancer activity through a specialized beta-globin TATA box. *Genes Dev* 6: 521-532, 1992.
36. Bagga R, Michalowski S, Sabnis R, Griffith JD and Emerson BM: HMG I/Y regulates long-range enhancer-dependent transcription on DNA and chromatin by changes in DNA topology. *Nucleic Acids Res* 28: 2541-2550, 2000.
37. Reeves R: HMGA proteins: flexibility finds a nuclear niche? *Biochem Cell Biol* 81: 185-195, 2003.
38. Yie J, Merika M, Munshi N, Chen G and Thanos D: The role of HMG I(Y) in the assembly and function of the IFN-beta enhanceosome. *EMBO J* 18: 3074-3089, 1999.
39. Carvajal IM, Baron RM and Perrella MA: High-mobility group-I/Y proteins: potential role in the pathophysiology of critical illnesses. *Crit Care Med* 30: S36-S42, 2002.
40. Reeves R: Molecular biology of HMGA proteins: hubs of nuclear function. *Gene* 277: 63-81, 2001.
41. Tessari MA, Gostissa M, Altamura S, *et al*: Transcriptional activation of the cyclin A gene by the architectural transcription factor HMGA2. *Mol Cell Biol* 23: 9104-9116, 2003.
42. Farnet CM and Bushman FD: HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes *in vitro*. *Cell* 88: 483-492, 1997.
43. Li L, Yoder K, Hansen MS, Olvera J, Miller MD and Bushman FD: Retroviral cDNA integration: stimulation by HMG I family proteins. *J Virol* 74: 10965-10974, 2000.
44. Li L, Farnet CM, Anderson WF and Bushman FD: Modulation of activity of Moloney murine leukemia virus preintegration complexes by host factors *in vitro*. *J Virol* 72: 2125-2131, 1998.
45. Hindmarsh P, Ridky T, Reeves R, Andrade M, Skalka AM and Leis J: HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration *in vitro*. *J Virol* 73: 2994-3003, 1999.
46. Hindmarsh P and Leis J: Retroviral DNA integration. *Microbiol Mol Biol Rev* 63: 836-843, 1999.
47. Van Maele B, Busschots K, Vandekerckhove L, Christ F and Debyser Z: Cellular co-factors of HIV-1 integration. *Trends Biochem Sci* 31: 98-105, 2006.
48. Leger H, Sock E, Renner K, Grummt F and Wegner M: Functional interaction between the POU domain protein Tst-1/Oct-6 and the high-mobility-group protein HMG-I/Y. *Mol Cell Biol* 15: 3738-3747, 1995.
49. French SW, Schmidt MC and Glorioso JC: Involvement of a high-mobility-group protein in the transcriptional activity of herpes simplex virus latency-active promoter 2. *Mol Cell Biol* 16: 5393-5399, 1996.
50. Bouallaga I, Teissier S, Yaniv M and Thierry F: HMG-I(Y) and the CBP/p300 coactivator are essential for human papillomavirus type 18 enhanceosome transcriptional activity. *Mol Cell Biol* 23: 2329-2340, 2003.
51. Bouallaga I, Massicard S, Yaniv M and Thierry F: An enhanceosome containing the Jun B/Fra-2 heterodimer and the HMG-I(Y) architectural protein controls HPV 18 transcription. *EMBO Rep* 1: 422-427, 2000.
52. Henderson A, Bunce M, Siddon N, Reeves R and Tremethick DJ: High-mobility-group protein I can modulate binding of transcription factors to the U5 region of the human immunodeficiency virus type 1 proviral promoter. *J Virol* 74: 10523-10534, 2000.
53. Li Y, Lu J and Prochownik EV: Dual role for SUMO E<sub>2</sub> conjugase Ubc9 in modulating the transforming and growth-promoting properties of the HMGA1b architectural transcription factor. *J Biol Chem* 282: 13363-13371, 2007.
54. Giancotti V, Bandiera A, Sindici C, Perissin L and Crane-Robinson C: Calcium-dependent ADP-ribosylation of high-mobility-group I (HMGI) proteins. *Biochem J* 317: 865-870, 1996.
55. Wang DZ, Ray P and Boothby M: Interleukin 4-inducible phosphorylation of HMG-I(Y) is inhibited by rapamycin. *J Biol Chem* 270: 22924-22932, 1995.
56. Banks GC, Li Y and Reeves R: Differential *in vivo* modifications of the HMGI(Y) non-histone chromatin proteins modulate nucleosome and DNA interactions. *Biochemistry* 39: 8333-8346, 2000.
57. Schwanbeck R, Manfioletti G and Wisniewski JR: Architecture of high mobility group protein I-C.DNA complex and its perturbation upon phosphorylation by Cdc2 kinase. *J Biol Chem* 275: 1793-1801, 2000.
58. Goodwin G: The high mobility group protein, HMGI-C. *Int J Biochem Cell Biol* 30: 761-766, 1998.
59. Di Agostino S, Fedele M, Chieffi P, Fusco A, Rossi P, Geremia R and Sette C: Phosphorylation of high-mobility group protein A2 by Nek2 kinase during the first meiotic division in mouse spermatocytes. *Mol Biol Cell* 15: 1224-1232, 2004.
60. Munshi N, Merika M, Yie J, Senger K, Chen G and Thanos D: Acetylation of HMG I(Y) by CBP turns off IFN beta expression by disrupting the enhanceosome. *Mol Cell* 2: 457-467, 1998.
61. Munshi N, Agaloti T, Lomvardas S, Merika M, Chen G and Thanos D: Coordination of a transcriptional switch by HMGI(Y) acetylation. *Science* 293: 1133-1136, 2001.
62. Zhang Q, Zhang K, Zou Y, Perna A and Wang Y: A quantitative study on the *in vitro* and *in vivo* acetylation of high mobility group A1 proteins. *J Am Soc Mass Spectrom* 18: 1569-1578, 2007.
63. Diana F, Sgarra R, Manfioletti G, *et al*: A link between apoptosis and degree of phosphorylation of high mobility group A1a protein in leukemic cells. *J Biol Chem* 276: 11354-11361, 2001.
64. Edberg DD, Adkins JN, Springer DL and Reeves R: Dynamic and differential *in vivo* modifications of the isoform HMGA1a and HMGA1b chromatin proteins. *J Biol Chem* 280: 8961-8973, 2005.
65. Edberg DD, Bruce JE, Siems WF and Reeves R: *In vivo* post-translational modifications of the high mobility group A1a proteins in breast cancer cells of differing metastatic potential. *Biochemistry* 43: 11500-11515, 2004.
66. Sgarra R, Lee J, Tessari MA, *et al*: The AT-hook of the chromatin architectural transcription factor high mobility group A1a is arginine-methylated by protein arginine methyltransferase 6. *J Biol Chem* 281: 3764-3772, 2006.
67. Zou Y, Webb K, Perna AD, Zhang Q, Clarke S and Wang Y: A mass spectrometric study on the *in vitro* methylation of HMGA1a and HMGA1b proteins by PRMTs: methylation specificity, the effect of binding to AT-rich duplex DNA, and the effect of C-terminal phosphorylation. *Biochemistry* 46: 7896-7906, 2007.
68. Baldassarre G, Fedele M, Battista S, *et al*: Onset of natural killer cell lymphomas in transgenic mice carrying a truncated HMGI-C gene by the chronic stimulation of the IL-2 and IL-15 pathway. *Proc Natl Acad Sci USA* 98: 7970-7975, 2001.
69. Fedele M, Pentimalli F, Baldassarre G, *et al*: Transgenic mice overexpressing the wild-type form of the HMGA1 gene develop mixed growth hormone/prolactin cell pituitary adenomas and natural killer cell lymphomas. *Oncogene* 24: 3427-3435, 2005.
70. Baldassarre G, Battista S, Belletti B, *et al*: Negative regulation of BRCA1 gene expression by HMGA1 proteins accounts for the reduced BRCA1 protein levels in sporadic breast carcinoma. *Mol Cell Biol* 23: 2225-2238, 2003.

71. Cmarik JL, Li Y, Ogram SA, Min H, Reeves R and Colburn NH: Tumor promoter induces high mobility group HMG-Y protein expression in transformation-sensitive but not -resistant cells. *Oncogene* 16: 3387-3396, 1998.
72. Reeves R, Edberg DD and Li Y: Architectural transcription factor HMGI(Y) promotes tumor progression and mesenchymal transition of human epithelial cells. *Mol Cell Biol* 21: 575-594, 2001.
73. Green MC: Catalog of mutant genes and polymorphic loci. In: Genetic Variants and Strains of the Laboratory Mouse. Lyon M and Searle A (eds.) Oxford University Press, pp12-403, 1989.
74. Zhou X, Benson KF, Ashar HR and Chada K: Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature* 376: 771-774, 1995.
75. Benson KF and Chada K: Mini-mouse: phenotypic characterization of a transgenic insertional mutant allelic to pygmy. *Genet Res* 64: 27-33, 1994.
76. Brants JR, Ayoubi TA, Chada K, Marchal K, van de Ven WJ and Petit MM: Differential regulation of the insulin-like growth factor II mRNA-binding protein genes by architectural transcription factor HMGA2. *FEBS Lett* 569: 277-283, 2004.
77. Cleyne I, Brants JR, Peeters K, *et al*: HMGA2 regulates transcription of the Imp2 gene via an intronic regulatory element in cooperation with nuclear factor-kappaB. *Mol Cancer Res* 5: 363-372, 2007.
78. Lanahan A, Williams JB, Sanders LK and Nathans D: Growth factor-induced delayed early response genes. *Mol Cell Biol* 12: 3919-3929, 1992.
79. Chiappetta G, Avantaggiato V, Visconti R, *et al*: High level expression of the HMGI (Y) gene during embryonic development. *Oncogene* 13: 2439-2446, 1996.
80. Hirning-Folz U, Wilda M, Rippe V, Bullerdiek J and Hameister H: The expression pattern of the Hmgic gene during development. *Genes Chromosomes Cancer* 23: 350-357, 1998.
81. Zhou X, Benson KF, Przybysz K, Liu J, Hou Y, Cherath L and Chada K: Genomic structure and expression of the murine Hmgic gene. *Nucleic Acids Res* 24: 4071-4077, 1996.
82. Rogalla P, Drechsler K, Frey G, Hennig Y, Helmke B, Bonk U and Bullerdiek J: HMGI-C expression patterns in human tissues. Implications for the genesis of frequent mesenchymal tumors. *Am J Pathol* 149: 775-779, 1996.
83. Vartiainen E, Palvimo J, Mahonen A, Linnala-Kankkunen A and Maenpaa PH: Selective decrease in low-Mr HMG proteins HMG I and HMG Y during differentiation of mouse teratocarcinoma cells. *FEBS Lett* 228: 45-48, 1988.
84. Giancotti V, Pani B, D'Andrea P, *et al*: Elevated levels of a specific class of nuclear phosphoproteins in cells transformed with v-ras and v-mos oncogenes and by cotransfection with c-myc and polyoma middle T genes. *EMBO J* 6: 1981-1987, 1987.
85. Melillo RM, Pierantoni GM, Scala S, *et al*: Critical role of the HMGI(Y) proteins in adipocyte cell growth and differentiation. *Mol Cell Biol* 21: 2485-2495, 2001.
86. Smas CM and Sul HS: Control of adipocyte differentiation. *Biochem J* 309: 697-710, 1995.
87. Ayoubi TA, Jansen E, Meulemans SM and van de Ven WJ: Regulation of HMGIC expression: an architectural transcription factor involved in growth control and development. *Oncogene* 18: 5076-5087, 1999.
88. Battista S, Pentimalli F, Baldassarre G, Fedele M, Fidanza V, Croce CM and Fusco A: Loss of Hmgal gene function affects embryonic stem cell lympho-hematopoietic differentiation. *FASEB J* 17: 1496-1498, 2003.
89. Battista S, Fedele M, Martinez Hoyos J, *et al*: High-mobility-group A1 (HMGA1) proteins down-regulate the expression of the recombination activating gene 2 (RAG2). *Biochem J* 389: 91-97, 2005.
90. Liu J, Schiltz JF, Ashar HR and Chada KK: Hmgal is required for normal sperm development. *Mol Reprod Dev* 66: 81-89, 2003.
91. Fedele M, Fidanza V, Battista S, *et al*: Haploinsufficiency of the Hmgal gene causes cardiac hypertrophy and myelo-lymphoproliferative disorders in mice. *Cancer Res* 66: 2536-2543, 2006.
92. Caron L, Bost F, Prot M, Hofman P and Binetruy B: A new role for the oncogenic high-mobility group A2 transcription factor in myogenesis of embryonic stem cells. *Oncogene* 24: 6281-6291, 2005.
93. Foti D, Chiefari E, Fedele M, *et al*: Lack of the architectural factor HMGA1 causes insulin resistance and diabetes in humans and mice. *Nat Med* 11: 765-773, 2005.
94. Shannon MF, Himes SR and Attema J: A role for the architectural transcription factors HMGI(Y) in cytokine gene transcription in T cells. *Immunol Cell Biol* 76: 461-466, 1998.
95. Liu F, Chau KY, Arlotta P and Ono SJ: The HMG I proteins: dynamic roles in gene activation, development, and tumorigenesis. *Immunol Res* 24: 13-29, 2001.
96. Chau KY, Munshi N, Keane-Myers A, *et al*: The architectural transcription factor high mobility group I(Y) participates in photoreceptor-specific gene expression. *J Neurosci* 20: 7317-7324, 2000.
97. Bussemakers MJ, van de Ven WJ, Debruyne FM and Schalken JA: Identification of high mobility group protein I(Y) as potential progression marker for prostate cancer by differential hybridization analysis. *Cancer Res* 51: 606-611, 1991.
98. Ram TG, Reeves R and Hosick HL: Elevated high mobility group-I(Y) gene expression is associated with progressive trans-formation of mouse mammary epithelial cells. *Cancer Res* 53: 2655-2660, 1993.
99. Tamimi Y, van der Poel HG, Karthaus HF, Debruyne FM and Schalken JA: A retrospective study of high mobility group protein I(Y) as progression marker for prostate cancer determined by *in situ* hybridization. *Br J Cancer* 74: 573-578, 1996.
100. Tamimi Y, van der Poel HG, Denyn MM, Umbas R, Karthaus HF, Debruyne FM and Schalken JA: Increased expression of high mobility group protein I(Y) in high grade prostatic cancer determined by *in situ* hybridization. *Cancer Res* 53: 5512-5516, 1993.
101. Kim DH, Park YS, Park CJ, *et al*: Expression of the HMGI(Y) gene in human colorectal cancer. *Int J Cancer* 84: 376-380, 1999.
102. Fedele M, Bandiera A, Chiappetta G, *et al*: Human colorectal carcinomas express high levels of high mobility group HMGI(Y) proteins. *Cancer Res* 56: 1896-1901, 1996.
103. Chiappetta G, Manfioletti G, Pentimalli F, *et al*: High mobility group HMGI(Y) protein expression in human colorectal hyperplastic and neoplastic diseases. *Int J Cancer* 91: 147-151, 2001.
104. Sarhadi VK, Wikman H, Salmenkivi K, *et al*: Increased expression of high mobility group A proteins in lung cancer. *J Pathol* 209: 206-212, 2006.
105. Flohr AM, Rogalla P, Bonk U, *et al*: High mobility group protein HMGA1 expression in breast cancer reveals a positive correlation with tumour grade. *Histol Histopathol* 18: 999-1004, 2003.
106. Chiappetta G, Botti G, Monaco M, *et al*: HMGA1 protein overexpression in human breast carcinomas: correlation with ErbB2 expression. *Clin Cancer Res* 10: 7637-7644, 2004.
107. Abe N, Watanabe T, Masaki T, *et al*: Pancreatic duct cell carcinomas express high levels of high mobility group I(Y) proteins. *Cancer Res* 60: 3117-3122, 2000.
108. Abe N, Watanabe T, Izumisato Y, *et al*: Diagnostic significance of high mobility group I(Y) protein expression in intraductal papillary mucinous tumors of the pancreas. *Pancreas* 25: 198-204, 2002.
109. Chiappetta G, Bandiera A, Berlingieri MT, *et al*: The expression of the high mobility group HMGI (Y) proteins correlates with the malignant phenotype of human thyroid neoplasias. *Oncogene* 10: 1307-1314, 1995.
110. Rogalla P, Drechsler K, Kazmierczak B, Rippe V, Bonk U and Bullerdiek J: Expression of HMGI-C, a member of the high mobility group protein family, in a subset of breast cancers: relationship to histologic grade. *Mol Carcinog* 19: 153-156, 1997.
111. Berner JM, Meza-Zepeda LA, Kools PF, *et al*: HMGIC, the gene for an architectural transcription factor, is amplified and rearranged in a subset of human sarcomas. *Oncogene* 14: 2935-2941, 1997.
112. Abe N, Watanabe T, Suzuki Y, *et al*: An increased high-mobility group A2 expression level is associated with malignant phenotype in pancreatic exocrine tissue. *Br J Cancer* 89: 2104-2109, 2003.
113. Miyazawa J, Mitoro A, Kawashiri S, Chada KK and Imai K: Expression of mesenchyme-specific gene HMGA2 in squamous cell carcinomas of the oral cavity. *Cancer Res* 64: 2024-2029, 2004.
114. Meyer B, Loeschke S, Schultze A, *et al*: HMGA2 overexpression in non-small cell lung cancer. *Mol Carcinog* 46: 503-511, 2007.

115. Wood LJ, Maher JF, Bunton TE and Resar LM: The oncogenic properties of the HMG-I gene family. *Cancer Res* 60: 4256-4261, 2000.
116. Wood LJ, Mukherjee M, Dolde CE, *et al*: HMG-I/Y, a new c-Myc target gene and potential oncogene. *Mol Cell Biol* 20: 5490-5502, 2000.
117. Fedele M, Battista S, Kenyon L, *et al*: Overexpression of the HMGA2 gene in transgenic mice leads to the onset of pituitary adenomas. *Oncogene* 21: 3190-3198, 2002.
118. Finelli P, Pierantoni GM, Giardino D, *et al*: The high mobility group A2 gene is amplified and overexpressed in human pro-lactinomas. *Cancer Res* 62: 2398-2405, 2002.
119. Fedele M, Pierantoni GM, Visone R and Fusco A: E2F1 activation is responsible for pituitary adenomas induced by HMGA2 gene overexpression. *Cell Div* 1: 17, 2006.
120. Fedele M, Visone R, De Martino I, *et al*: HMGA2 induces pituitary tumorigenesis by enhancing E2F1 activity. *Cancer Cell* 9: 459-471, 2006.
121. Muller H, Bracken AP, Vernell R, *et al*: E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev* 15: 267-285, 2001.
122. Xu Y, Sumter TF, Bhattacharya R, *et al*: The HMG-I oncogene causes highly penetrant, aggressive lymphoid malignancy in transgenic mice and is overexpressed in human leukemia. *Cancer Res* 64: 3371-3375, 2004.
123. Pierantoni GM, Agosti V, Fedele M, *et al*: High-mobility group A1 proteins are overexpressed in human leukaemias. *Biochem J* 372: 145-150, 2003.
124. Tesfaye A, Di Cello F, Hillion J, *et al*: The high-mobility group A1 gene up-regulates cyclooxygenase 2 expression in uterine tumorigenesis. *Cancer Res* 67: 3998-4004, 2007.
125. Schoenmakers EF, Wanschura S, Mols R, Bullerdiek J, van den Berghe H and van de Ven WJ: Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. *Nat Genet* 10: 436-444, 1995.
126. Ashar HR, Fejzo MS, Tkachenko A, *et al*: Disruption of the architectural factor HMGI-C: DNA-binding AT hook motifs fused in lipomas to distinct transcriptional regulatory domains. *Cell* 82: 57-65, 1995.
127. Geurts JM, Schoenmakers EF and van de Ven WJ: Molecular characterization of a complex chromosomal rearrangement in a pleomorphic salivary gland adenoma involving the 3'-UTR of HMGIC. *Cancer Genet Cytogenet* 95: 198-205, 1997.
128. Mine N, Kurose K, Nagai H, *et al*: Gene fusion involving HMGIC is a frequent aberration in uterine leiomyomas. *J Hum Genet* 46: 408-412, 2001.
129. Van de Ven WJ: Genetic basis of uterine leiomyoma: involvement of high mobility group protein genes. *Eur J Obstet Gynecol Reprod Biol* 81: 289-293, 1998.
130. Kazmierczak B, Dal Cin P, Wanschura S, Borrmann L, Fusco A, van den Berghe H and Bullerdiek J: HMGIY is the target of 6p21.3 rearrangements in various benign mesenchymal tumors. *Genes Chromosomes Cancer* 23: 279-285, 1998.
131. Nucci MR, Weremowicz S, Neskey DM, Sornerberger K, Tallini G, Morton CC and Quade BJ: Chromosomal translocation t(8;12) induces aberrant HMGIC expression in aggressive angiomyxoma of the vulva. *Genes Chromosomes Cancer* 32: 172-176, 2001.
132. Kazmierczak B, Meyer-Bolte K, Tran KH, *et al*: A high frequency of tumors with rearrangements of genes of the HMGI(Y) family in a series of 191 pulmonary chondroid hamartomas. *Genes Chromosomes Cancer* 26: 125-133, 1999.
133. Dal Cin P, Wanschura S, Kazmierczak B, *et al*: Amplification and expression of the HMGIC gene in a benign endometrial polyp. *Genes Chromosomes Cancer* 22: 95-99, 1998.
134. Bol S, Wanschura S, Thode B, Deichert U, van de Ven WJ, Bartnitzke S and Bullerdiek J: An endometrial polyp with a rearrangement of HMGI-C underlying a complex cytogenetic rearrangement involving chromosomes 2 and 12. *Cancer Genet Cytogenet* 90: 88-90, 1996.
135. Hennig Y, Wanschura S, Deichert U, Bartnitzke S and Bullerdiek J: Rearrangements of the high mobility group protein family genes and the molecular genetic origin of uterine leiomyomas and endometrial polyps. *Mol Hum Reprod* 2: 277-283, 1996.
136. Petit MM, Mols R, Schoenmakers EF, Mandahl N and van de Ven WJ: LPP, the preferred fusion partner gene of HMGIC in lipomas, is a novel member of the LIM protein gene family. *Genomics* 36: 118-129, 1996.
137. Schoenmakers EF, Huysmans C and van de Ven WJ: Allelic knockout of novel splice variants of human recombination repair gene RAD51B in t(12;14) uterine leiomyomas. *Cancer Res* 59: 19-23, 1999.
138. Rogalla P, Kazmierczak B, Meyer-Bolte K, Tran KH and Bullerdiek J: The t(3;12)(q27;q14-q15) with underlying HMGIC-LPP fusion is not determining an adipocytic phenotype. *Genes Chromosomes Cancer* 22: 100-104, 1998.
139. Kazmierczak B, Hennig Y, Wanschura S, Rogalla P, Bartnitzke S, van de Ven W and Bullerdiek J: Description of a novel fusion transcript between HMGI-C, a gene encoding for a member of the high mobility group proteins, and the mitochondrial aldehyde dehydrogenase gene. *Cancer Res* 55: 6038-6039, 1995.
140. Kools PF and van de Ven WJ: Amplification of a rearranged form of the high-mobility group protein gene HMGIC in OsA-CI osteosarcoma cells. *Cancer Genet Cytogenet* 91: 1-7, 1996.
141. Schoenberg Fejzo M, Ashar HR, Krauter KS, *et al*: Translocation breakpoints upstream of the HMGIC gene in uterine leiomyomata suggest dysregulation of this gene by a mechanism different from that in lipomas. *Genes Chromosomes Cancer* 17: 1-6, 1996.
142. Quade BJ, Weremowicz S, Neskey DM, Vanni R, Ladd C, Dal Cin P and Morton CC: Fusion transcripts involving HMGA2 are not a common molecular mechanism in uterine leiomyomata with rearrangements in 12q15. *Cancer Res* 63: 1351-1358, 2003.
143. Pasquinelli AE, Reinhart BJ, Slack F, *et al*: Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408: 86-89, 2000.
144. Schulman BR, Esquela-Kerscher A and Slack FJ: Reciprocal expression of lin-41 and the microRNAs let-7 and mir-125 during mouse embryogenesis. *Dev Dyn* 234: 1046-1054, 2005.
145. Tallini G and Dal Cin P: HMGI(Y) and HMGI-C dysregulation: a common occurrence in human tumors. *Adv Anat Pathol* 6: 237-246, 1999.
146. Fedele M, Berlingieri MT, Scala S, *et al*: Truncated and chimeric HMGI-C genes induce neoplastic transformation of NIH3T3 murine fibroblasts. *Oncogene* 17: 413-418, 1998.
147. Crombez KR, Vanoirbeek EM, van de Ven WJ and Petit MM: Transactivation functions of the tumor-specific HMGA2/LPP fusion protein are augmented by wild-type HMGA2. *Mol Cancer Res* 3: 63-70, 2005.
148. Battista S, Fidanza V, Fedele M, *et al*: The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis. *Cancer Res* 59: 4793-4797, 1999.
149. Arlotta P, Tai AK, Manioletti G, Clifford C, Jay G and Ono SJ: Transgenic mice expressing a truncated form of the high mobility group I-C protein develop adiposity and an abnormally high prevalence of lipomas. *J Biol Chem* 275: 14394-14400, 2000.
150. Ligon AH, Moore SD, Parisi MA, *et al*: Constitutional rearrangement of the architectural factor HMGA2: a novel human phenotype including overgrowth and lipomas. *Am J Hum Genet* 76: 340-348, 2005.
151. Tallini G, Dal Cin P, Rhoden KJ, *et al*: Expression of HMGI-C and HMGI(Y) in ordinary lipoma and atypical lipomatous tumors: immunohistochemical reactivity correlates with karyotypic alterations. *Am J Pathol* 151: 37-43, 1997.
152. Kazmierczak B, Dal Cin P, Sciort R, van den Berghe H and Bullerdiek J: Inflammatory myofibroblastic tumor with HMGIC rearrangement. *Cancer Genet Cytogenet* 112: 156-160, 1999.
153. Kazmierczak B, Wanschura S, Rommel B, Bartnitzke S and Bullerdiek J: Ten pulmonary chondroid hamartomas with chromosome 6p21 breakpoints within the HMG-I(Y) gene or its immediate surroundings. *J Natl Cancer Inst* 88: 1234-1236, 1996.
154. Xiao S, Lux ML, Reeves R, Hudson TJ and Fletcher JA: HMGI(Y) activation by chromosome 6p21 rearrangements in multilineage mesenchymal cells from pulmonary hamartoma. *Am J Pathol* 150: 901-910, 1997.
155. Wada T, Takei K, Kudo M, *et al*: Characterization of immune function and analysis of RAG gene mutations in Omenn syndrome and related disorders. *Clin Exp Immunol* 119: 148-155, 2000.
156. Abe T, Tsuge I, Kamachi Y, *et al*: Evidence for defects in V(D)J rearrangements in patients with severe combined immunodeficiency. *J Immunol* 152: 5504-5513, 1994.

157. Foti D, Iuliano R, Chiefari E and Brunetti A: A nucleoprotein complex containing Sp1, C/EBP beta, and HMGI-Y controls human insulin receptor gene transcription. *Mol Cell Biol* 23: 2720-2732, 2003.
158. Brunetti A, Manfioletti G, Chiefari E, Goldfine ID and Foti D: Transcriptional regulation of human insulin receptor gene by the high-mobility group protein HMGI(Y). *FASEB J* 15: 492-500, 2001.
159. Giannini G, Cerignoli F, Mellone M, Massimi I, Ambrosi C, Rinaldi C and Gulino A: Molecular mechanism of HMGA1 deregulation in human neuroblastoma. *Cancer Lett* 228: 97-104, 2005.
160. Ogram SA and Reeves R: Differential regulation of a multi-promoter gene. Selective 12-O-tetradecanoylphorbol-13-acetate induction of a single transcription start site in the HMG-I/Y gene. *J Biol Chem* 270: 14235-14242, 1995.
161. Giannini G, Di Marcotullio L, Ristori E, *et al*: HMGI(Y) and HMGI-C genes are expressed in neuroblastoma cell lines and tumors and affect retinoic acid responsiveness. *Cancer Res* 59: 2484-2492, 1999.
162. Giannini G, Cerignoli F, Mellone M, *et al*: High mobility group A1 is a molecular target for MYCN in human neuroblastoma. *Cancer Res* 65: 8308-8316, 2005.
163. Hommura F, Katabami M, Leaner VD, Donniger H, Sumter TF, Resar LM and Birrer MJ: HMG-I/Y is a c-Jun/activator protein-1 target gene and is necessary for c-Jun-induced anchorage-independent growth in Rat1a cells. *Mol Cancer Res* 2: 305-314, 2004.
164. Cleynen I, Huysmans C, Sasazuki T, Shirasawa S, van de Ven W and Peeters K: Transcriptional control of the human high mobility group A1 gene: basal and oncogenic Ras-regulated expression. *Cancer Res* 67: 4620-4629, 2007.
165. Borrmann L, Wilkening S and Bullerdiek J: The expression of HMGA genes is regulated by their 3'UTR. *Oncogene* 20: 4537-4541, 2001.
166. Lee YS and Dutta A: The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev* 21: 1025-1030, 2007.
167. Mayr C, Hemann MT and Bartel DP: Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* 315: 1576-1579, 2007.
168. Tkachenko A, Ashar HR, Meloni AM, Sandberg AA and Chada KK: Misexpression of disrupted HMGI architectural factors activates alternative pathways of tumorigenesis. *Cancer Res* 57: 2276-2280, 1997.
169. Brants JR: The Insulin-like growth factor II mRNA-binding protein 2 gene, a downstream target of the architectural transcription factor HMGA2. Thesis manuscript, Center for Human Genetics. KU Leuven, p147, 2005.