Cloning and knock-out of the mouse gene coding for the high mobility group 2 protein (HMG2)

Thesis

UNRESTRICTED

CLONING AND KNOCK-OUT OF THE MOUSE GENE CODING FOR THE HIGH MOBILITY GROUP 2 PROTEIN (HMG2)

Lorenza Ronfani

A thesis submitted in partial fulfilment of the requirement of the Open University for the degree of Doctor of Philosophy

January 2000

DIBIT
Department of Biological and Technical Research
Istituto Scientifico Ospedale San Raffaele
Milan, Italy

DATE OF SUBMISSION : 31 JANUARY 2000
DATE OF AWARD : 17 APRIL 2000

CONTENTS

ABSTRACT

High Mobility Group proteins 1 and 2 are highly conserved nuclear proteins ubiquitously expressed in higher eukaryotic cells. HMG1 and HMG2 each contain two similar DNA binding domains, called HMG-boxes, and an acidic tail. HMG1 and HMG2 are able to facilitate the binding to DNA of several transcription factors (Steroid Hormone Receptors, HOX and OCT proteins) and they are involved in V(D)J recombination.

hmg1 -/- mice, produced in our lab by conventional knock-out, are born and die within the first day of life. The cause of death is a severe hypoglycaemia: the level of glucose is low in the blood of mutant mice, while abundant glycogen is still present in their liver. Glucocorticoid-dependent gene expression is impaired in -/- mice. Surprisingly, HMG1 is not essential for the life of the cell, as it might be suggested by the abundance of the protein and its evolutionary conservation. We then suggested that another protein can substitute HMG1 in its function in the cell. The best candidate is HMG2 because of its remarkable similarity with HMG1. Therefore, I decided to investigate the function of HMG2 by gene targeting, and I planned the work so as to investigate similarities and differences between HMG1 and HMG2.

The mouse *hmg2* gene was cloned, characterised and mapped on the centromeric region of chromosome 8. Further, functional analysis were done. By transient transfection assays, I demonstrated that HMG2 is very similar to HMG1 in the transcriptional activation mediated by HOX proteins. Using the same assay, I demonstrated that differences in the length of the acidic tail of HMG1 and HMG2 determine no functional differences. Moreover, I showed that HMG2, as HMG1, is not stably associated to chromatin. These data, together with previously published data,

suggest that HMG2 and HMG1 play very similar functions in the cell. Nevertheless, a difference between the two proteins was found in their expression pattern. HMG2 is absent in adult liver and brain, while is very abundant in adult testis, spleen, and thymus. On the contrary, HMG1 is ubiquitously expressed. *In situ* hybridisation revealed that in the testis HMG2 has a specific distribution: it is absent in spermatogonia and in late spermatids and spermatozoa, while is very abundant in spermatocytes and round spermatids. The peculiar protein distribution during spermatogenesis correlates with the phenotype we found in *hmg2* $^{-/-}$ mice.

hmg2 $^{-/-}$ mice had no obvious phenotypic differences from wild type. No immunological, homeotic, and chromatin defects were found. However, -/- male mice are partially sterile. After six months of continuous breeding studies some -/- mice gave no litters, some others gave litters of reduced size while others gave normal litters. Histological analysis of the testis in completely sterile mice revealed the presence of a significant number of dysmorphogenic seminiferous tubules, similar to that described in knock-out mice for estrogen receptor and in mice overexpressing ABP (Androgen Binding Protein). Degenerated tubules were also found in fertile -/- mice, but the number was drastically reduced. Moreover, electron microscope analysis revealed severe defects in elongating spermatids. Spermatozoa are produced, but are mostly immobile. Nevertheless, they are able to fertilise eggs.

The testes of many knock-outs were found in the inguinal ring instead of in the scrotum. The correct position of testes is regulated by androgens. The correct function of androgen receptor was tested in mouse embryonic fibroblasts, but no large difference in the response to testosterone was found in this kind of cells.

Thus, HMG1 and HMG2 play a very similar role in the cell and it is highly probable that they can substitute each other to a large extent. HMG1 expression is high and ubiquitous, whereas HMG2 expression is limited to a subset of tissues and organs. In the testes, the absence of HMG2 leads to a reduction in the number of viable spermatozoa, suggesting that during male meiosis a high level of HMGs is required.

1 INTRODUCTION

1.1 HMG proteins (High Mobility Group)

HMG proteins are a heterogeneous group of very abundant non-histonic proteins present in the nucleus of all eukaryotic cells. These proteins are characterised by high mobility in a denaturing SDS acrylamide gel, hence the acronym High Mobility Group (Bustin et al., 1990).

At the beginning of the seventies, the HMG proteins were classified by their biochemical properties: they can be extracted from chromatin with 0,35 M NaCl, are soluble in 2% tricloracetic acid and 5% perchloric acid and have a molecular weight lower than 30 KDa (Bustin et al., 1990).

The HMG proteins consist of three families: HMG-I/Y, HMG-14/17 and HMG1/2.

HMG-I and its isoform HMG-Y are small proteins of 11.9 KDa and 10.6 KDa, respectively. These proteins are present in proliferative or undifferentiated cells and in neoplastic tissues. They preferentially bind AT-rich DNA sequences in heterochromatin. Moreover, these proteins interact with AT-rich sequences of promoter regions and enhancers (Thanos and Maniatis, 1992).

HMG-14/17 are small proteins of 10 KDa, are highly charged and contain basic amino acids in the N-terminal region and acidic amino acids in the C-terminal region. These proteins have high affinity for the nucleosome and they bind the core in a stechiometric ratio of two molecules per nucleosome in actively transcribed chromatin (Crippa et al., 1992).

HMG1 and HMG2 are proteins with a molecular weight of 25 KDa. This family of proteins does not have any homology with HMG-14/17 and HMG-I/Y. HMG1 is the most abundant non-histonic protein, is present in the nucleus of all vertebrates and its amino acid sequence is extremely conserved in all species. HMG2 is highly related to

HMG1 but is present in the nucleus with a lower abundance. At the start of my work, the study of these proteins suggested that they perform a number of functions in the cell, but none had been characterised in great detail. As I will describe later, the most obvious characteristic of these proteins is the presence of a particular DNA binding domain, called HMG-box.

1.2 The HMG-box domain and HMG-box proteins

Proteins that interact with DNA play a crucial role in the cell. They are involved in several biological processes, such as DNA replication, transcription of active genes, and repair of damaged DNA. These proteins interact with DNA by means of domains that belong to distinct families: for example, "zinc finger", "helix-turn-helix", "helix-loop-helix", "leucine zipper", etc.

The existence of a new structural DNA binding motif, conserved during evolution and present in several eukaryotic proteins, was suggested by comparison of the amino acid sequence between hUBF, a transcription factor for human RNA polymerase I, and HMG1 (Jantzen et al., 1990). Then, several eukaryotic proteins were found containing this DNA binding domain, called HMG-box. These proteins were classified in three subgroups depending on their sequence selectivity.

The first group consists of HMG1 and HMG2. They contain two HMG-box domains that recognise a specific DNA structure (see Chapter 1.3).

The second group consists of general transcription factors of mitochondrial and ribosomal RNA polymerases, such as UBF and mtTF1. They do not bind a specific consensus sequence, but recognise partially specific sequences in the regions upstream of the transcription start sites.

The third group of proteins are tissue-specific transcription factors that are able to recognise specific DNA sequences in the promoter region of different genes. This group comprises a number of fungal proteins involved in mating-type expression (Staben and

Yanofski, 1990; Sugimoto et al., 1991), the mammalian testis determining factor SRY (Gubbay et al., 1990; Sinclair et al., 1990), the protein product of *Sox* genes and a set of lymphoid specific enhancer binding factors (Travis et al., 1991; van de Wetering et al., 1991).

These proteins have in common a marked effect on DNA structure and an affinity for distorted DNA structure, such as four-way junctions and cisplatinum adducts (Bianchi et al., 1989; Pil et al., 1993; Pil and Lippard, 1992).

Sequence analysis of the proteins indicates that the HMG-box is minimally a stretch of about 70 amino acids, with a net positive charge and rich in aromatic residues and prolines. No amino acid is absolutely conserved, only three residues have conservative substitutions in all known HMG-boxes: an aromatic residue in position 11, a tryptophan (rarely a tyrosine) in position 41 and a tyrosine (rarely a tryptophan or a phenylalanine) in position 52, while 21 other residues are not necessarily conserved (Fig.1). The sequence variation between HMG-boxes is paralleled by the diversity of their presumed biochemical functions.

The NMR structures of HMG-boxes of non-sequence-selective class exhibited the same general fold with minor differences, in which three α-helical segments formed an L-shaped structure stabilised by hydrophobic core. The shorter arm of the "L" consists of helices I and II and the longer arm is composed of the extended N-terminal section packed against the C-terminal region, helix III; the angle between the two arms is around 80 degrees (Read et al., 1993; Read et al., 1994; Weir et al., 1993) (Fig.2).

Recently, the X-ray structure of the rat HMG-boxA domain complexed to cisplatin-modified DNA was published (Ohndorf et al., 1999). This is the first resolution of any complex between a structure-specific HMG domain and a DNA substrate. Cisplatinum is a drug that damages DNA, creating intrastrand crosslinks between adjacent purines (Pil and Lippard, 1992). These crosslinks bend and unwind the duplex, and the altered structure is recognised by HMG1. In the X-ray structure shown in Fig.3, the

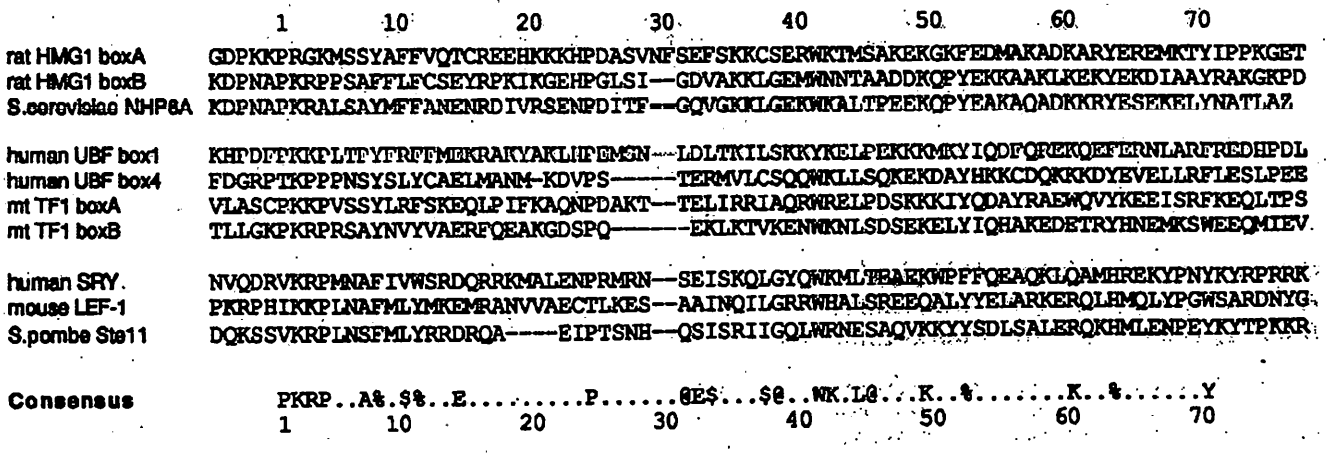


Fig.1.Alignment of HMG-boxes. The first group contains HMG-boxes from chromatin proteins. The second group of HMG-boxes are from general transcription factors for RNA polymerase I (human UBF, with four boxes) and mitochondrial RNA polymerases (human mtTFI, with two boxes). The third group are from tissue-specific transcription factors, as LEF1 and SRY, that recognise specific DNA sequences. Dashes indicate gaps in the alignment; Z indicates a stop codon in the gene. The consensus sequence for the HMG-box motif was obtained from 21 protein sequences; one-letter symbol indicate amino acids present in 50%-90% of the sequences; conservative substitutions (at least 75% of the occurrences at a particular position) are indicated as follows: @ for proline, alanine, glycine, serine, and threonine; % for tryptophan, phenylalanine and tyrosine; $ for methionine, valine, leucine and isoleucine. The numbering system starts from the first conserved proline and is based on HMG1-boxA.

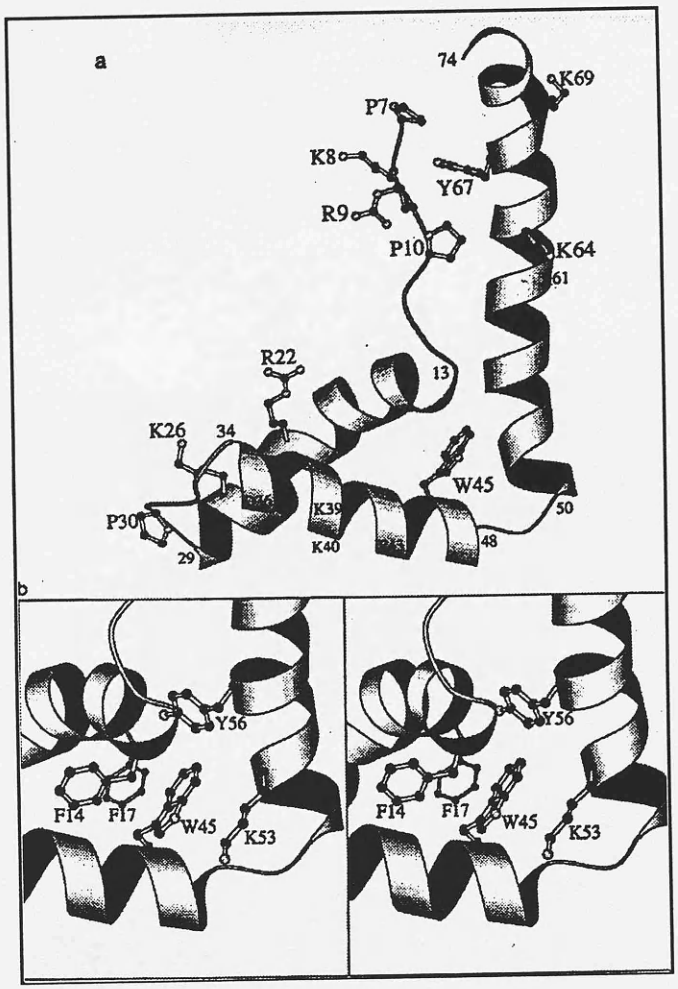


Fig.2. Schematic representations of one of the 30 structures of the B-domain HMG-box, showing the location of some of the conserved residues. (a) The three a-helical segments form an L-shaped structure. (b) A stereo view of the cluster of conserved residues at the junction of the two arms of the structure.

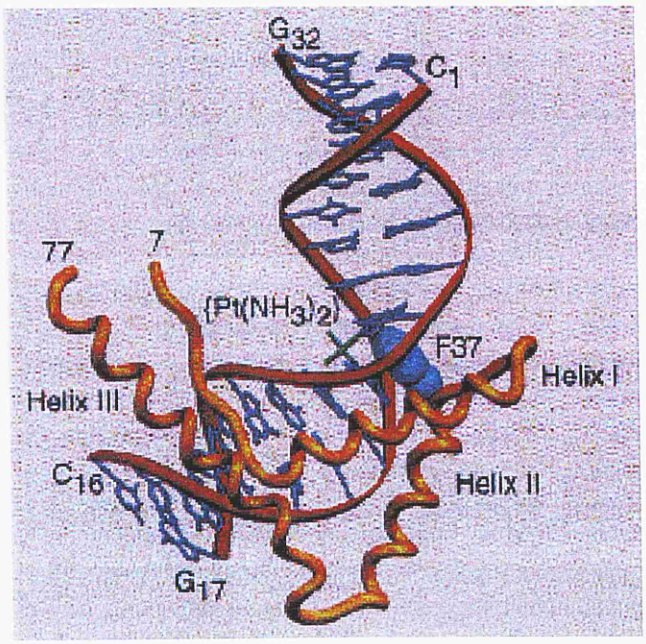


Fig.3. Structure of the complex between cisplatin-modified DNA and HMG-boxA. The protein backbone is shown in yellow, the intercalating Phe 37 residue as van der Waals spheres, the DNA in red and blue with the cisplatinum intrastrand adduct in green. Numbers indicate the first (N terminus) and last (C terminus) ordered residues in the crystal structure.

DNA is strongly kinked at a hydrophobic notch created at the platinum-DNA crosslink, and protein binding extends exclusively to the 3' side of the platinated strand. A phenylalanine residue at position 37 intercalates into the hydrophobic cleft at the cisplatin crosslink. The substitution of Phe37 with an alanine dramatically reduces the binding of the domain. Helix I of HMG-boxA interacts with the sugar-phosphate backbone of the unmodified strand, whereas helix II contacts the backbone of the platinated strand. The N-termini of the two helices fit in the minor groove.

The NMR structure of the HMG-box of SRY and of its DNA binding site, the promoter of the *MIS* gene (Müllerian Inhibiting Substance), shows that the overall fold is closely similar to the HMG-box of HMG1, but the angle of the "L" is more obtuse. The concave surface of the protein is perfectly located in the minor groove of the DNA where it covers a large surface area. To achieve this the DNA is distorted by 70 to 80 degrees (Werner et al., 1995).

The amino acid comparison between the HMGs that recognise DNA in a sequence specific manner (hSRY-box) and the HMGs that recognise specific DNA structure (HMG1-boxB) reveals differences of the residues that contact the DNA. In fact, the N- and C-terminal regions of hSRY-box contain hydrophobic residues (Val5, Tyr69 and Tyr72) that determine the formation of a stable structure able to recognise the consensus DNA sequence. In contrast, the N and C-terminal regions of HMG1-boxB do not contain hydrophobic residues and are highly disorganised. The lack of a stable structure in the N- and C-terminal region in non-sequence selective proteins could determine the inability to recognise specific DNA sequences. The amino acid at position 37, could also play an important role to discriminate sequence- or structure specific recognition of DNA. In fact, all non-sequence-selective class HMGs, including HMG1/2, HMG-D and NHP6A, have a hydrophobic residue at this position that can serve as a bending wedge, but which does not seem to discriminate on the basis of sequence. The HMG-boxes of SRY and LEF1

have a polar residue at position 37, which would allow sequence-specific hydrogen-bond formation.

1.3 The HMG1/2 family

HMG1 and HMG2 have a very similar structure. They have two DNA binding domains of the HMG-box type, designated A and B, and an acidic tail. HMG-boxA and B have a high concentration of basic amino acids, while the acidic tail is composed only by aspartic and glutammic acid and probably lacks a defined structure.

A new member of the HMG1/2 family, HMG4, has been recently cloned in our laboratory (Vaccari et al., 1998). HMG1, HMG2 and HMG4 are very similar to each other and are extremely conserved in vertebrates. The homology of murine proteins is 86% between HMG1 and 2, 89% between HMG1 and 4 and 85% between HMG2 and 4. The acidic tail is longer in HMG1 (30 aa), intermediate in HMG2 (25 aa) and shorter in HMG4 (20 aa). The high homology between the proteins of this family suggests that their genes arose from an ancestral duplication of a common precursor, as shown by the phylogenetic tree in Fig.4 (Vaccari et al., 1998). Recently, a lamprey cDNA was cloned, that codes for a protein highly homologous to HMG1 and 2, designated *Lfl*HMG1 (Sharman et al., 1997). The lamprey is a representative of the most ancestral class of vertebrate, the jawless. A philogenetic analysis showed that *Lfl*HMG1 is descended from a gene ancestral to mammalian HMG1 and HMG2. This implies that there was a duplication event that occurred after the divergence of the jawed and jawless fishes. Moreover, since pairs of proteins homologous to mammalian HMG1 and HMG2 were found in trout (*Teleostei*) and in Xenopus (*Amphibians*), the duplication of HMG1/2 originated before protostomes and deuterostomes diverged (Sharman et al., 1997). The relative introns position is conserved not only in the HMG1 family genes, but also in the *Drosophila DSP1* gene, which shows high homology to mammalian HMGs, and contains introns exactly in the same place as the mammalian genes (Canaple et al., 1997). It is

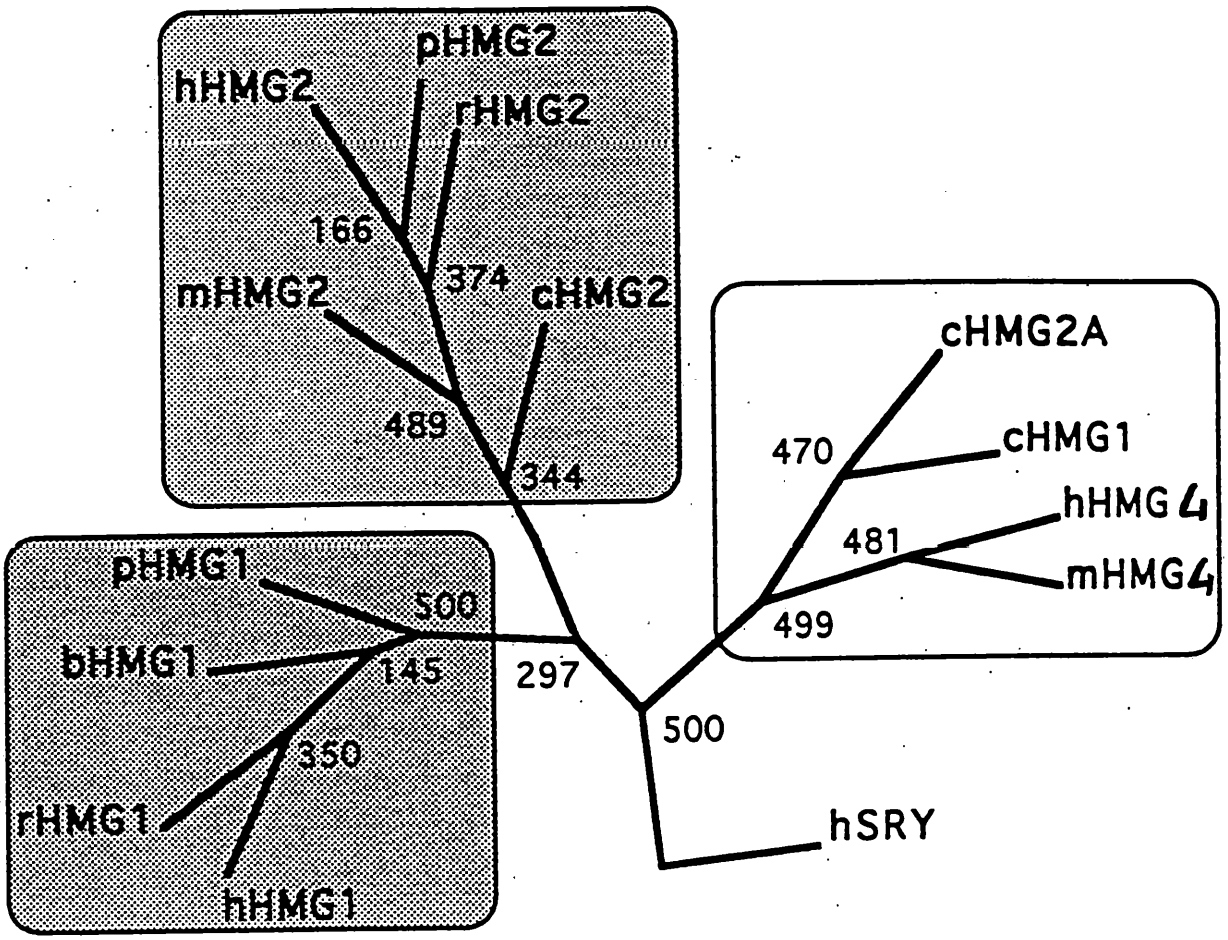


Fig.4 Phylogenetic tree of the HMG1/2 family (from Vaccari et al., 1998). The human SRY protein was used as outgroup. Three subgroups were found: the group that contains mammalian HMG1 proteins, the group that contains mammalian and chick HMG2 proteins, and the group that contains mouse and human HMG4 together with chick HMG1 and 2A. As shown by the tree HMG1, HMG2 and HMG4 subgroups derive from a common ancestral protein.

therefore likely that the internal duplication that originated the two boxes predates the arthropod/chordate separation (Canaple et al., 1997; Sharman et al., 1997).

Historically, HMG1 is the most studied of the HMG1/2 family. The abundance and the high degree of conservation between different species and during evolution suggests a fundamental role for this protein in the cell. Many years ago, it was shown by *in vitro* assays that HMG1 is able to bind with high affinity cruciform DNA (Bianchi et al., 1989). The binding is not sequence specific: in fact, HMG1 forms a specific complex with cruciform DNA, but not with linear DNA, single or double stranded, containing the same sequence present in cruciform DNA. Moreover, HMG1 specifically binds bulged DNA, where the presence of supernumerary bases on one of the two helices generates a strong distortion in the linear DNA molecule (Pontiggia and Bianchi, unpublished data). HMG1 bends the double helix significantly on binding through the minor groove (Pil et al., 1993) and binds with high affinity to DNA that is already sharply bent, such as linker DNA at the entry and exit of nucleosomes (Falciola et al., 1997; Travers et al., 1994; Ura and Wolffe, 1996).

Contradictory results have been obtained in the attempts to elucidate the role, if any, played by HMG1 and 2 proteins in the regulation of chromatin structure. An important role was suggested in DNA replication, chromatin assembly and disassembly (Bonne-Andrea et al., 1984; Travers et al., 1994; Waga et al., 1990), and transcription (Singh and Dixon, 1990; Tremethick and Molloy, 1986; 1988); however, none of these hypothesis has been confirmed unequivocally. Further, a role similar to histone H1 in the organisation and/or maintenance of chromatin structure was proposed for HMG1. HMG1 and HMG2 were described to be associated with the linker DNA of chromatin (Einck and Bustin, 1985; Goodwin et al., 1977; Scovell et al., 1987). Some authors reported the *in vitro* interaction between HMG1/2 and the histones (Bernues et al., 1983; Carballo et al., 1983). Xenopus HMG1 binds to nucleosomes *in vitro* in the same way of histone H1 and appears to replace histone H1 during early embryogenesis (Dimitrov et al., 1993; 1994;

Nightingale et al., 1996). Likewise, HMG-D, a Drosophila homologue of HMG1, associates with condensed chromatin during embryonal development and is gradually replaced by histone H1 after the mid-blastula transition (Ner and Travers, 1994).

Our lab investigated whether HMG1 plays overlapping or complementary roles with histone H1 (Falciola et al., 1997). We showed that in mammalian cells in culture, HMG1 and histone H1 differ widely in their ability to associate to chromatin. In fact, HMG1 binds *in vitro* to reconstituted core nucleosomes but is not stably associated with chromatin in live cells. During metaphase, HMG1, like many transcription factors, is detached from condensed chromosomes and diffuses from the nucleus to the cytoplasm. On the contrary, histone H1 remains bound to mitotic chromosomes. Moreover, HMG1 diffuses out of the nucleus during interphase after permeabilisation of the nuclear membranes with detergents, whereas histone H1 remains associated with chromatin. Recently, our lab also examined the association of HMG1 with chromatin in very early mouse development, a period when histone H1 is in very low abundance (Clarke et al., 1992). Throughout pre-implantation development, HMG1 protein always localised to the nucleus of interphase cells, but diffused to the cytoplasm and was never associated to condensed chromosomes during mitosis (Spada et al., 1998). This occurred even in the absence of significant levels of histone H1, indicating that the exclusion of HMG1 from condensed chromosomes was not due to competition. These results indicate that HMG1 does not substitute histone H1 in the organisation of chromatin during murine embryogenesis, contrary to what proposed for *Xenopus* and *Drosophila* embryogenesis (Ner and Travers, 1994; Nightingale et al., 1996; Ura and Wolffe, 1996). The substitution of HMGD for histone H1 in *Xenopus* and *Drosophila* embryos may be a particular characteristic of the more extended and less stable, rapidly replicating chromatin found in these species prior to the mid-blastula transition.

1.4 HMG1 and HMG2 as architectural regulators of transcription

Recently, HMG1 and HMG2 were proposed to act as co-factors of transcriptional activation.

Steroid hormone receptors exhibit weak binding affinity for their optimal palindromic hormone response elements (HREs). HMG1 and HMG2 were purified as factors able to increase the binding between steroid hormone receptors and their HREs. The two proteins make protein-protein interaction with steroid hormone receptors via their HMG-boxes and DNA binding domain (DBD), respectively, and do not bind the HRE elements (Boonyaratanakornkit et al., 1998; Oñate et al., 1994). By transient transfection assays, coexpression of HMG1 or HMG2 and steroid hormone receptors determines the enhancement of the transcriptional activation of a reporter gene under the control of HREs of 7 to 10 fold (Boonyaratanakornkit et al., 1998). The specificity of action of HMG1 and HMG2 was demonstrated with all the steroid receptors, on the contrary they have no effect on DNA binding by several nonsteroid nuclear receptors, including retinoic acid receptor (RAR), retinoic X receptor (RXR), and vitamin D receptor (VDR) (Boonyaratanakornkit et al., 1998).

A similar role was suggested for HMG2, in another context. In an expression screening for proteins that interact with the Octamer Transcription Factor Oct2, HMG2 was found (Zwilling et al., 1995). The analysis of the interaction revealed that the two proteins interact via the HMG domain and POU homeodomain, respectively. Oct2 is a lymphoid specific transcription factor, essential for the B cell-specific activities of immunoglobulin promoters (Bergman et al., 1984; Jenuwein and Grosschedl, 1991; Wirth et al., 1987). Also in this case the interaction between HMG2 and Oct2 determines an increase in sequence specific DNA interaction *in vitro*, and an enhancement in the transcriptional activation *in vivo*. HMG2 also interacts with two other octamer transcription factors, Oct1 and Oct6. Also in this case HMG2 appears to be a general co-

factor in the transcriptional activation mediated by octamer transcription factors (Zwilling et al., 1995).

Another group of proteins whose binding to DNA is enhanced by HMG1 are the HOX proteins. *Hox* genes are homologous to the Drosophila homeotic selector (*HOM*) genes, and encode sequence-specific transcription factors controlling the organisation of the body plan during development (reviewed by Krumlauf, 1994). The DNA binding domain of *Hox* genes, called homeodomain, is composed of three α-helices, and is structurally related to the "helix-turn-helix" motif. The homeodomain and flanking amino acids are involved in the DNA sequence recognition and binding *in vitro*. *In vivo* assays clearly demonstrated a specific recognition of target genes (Krumlauf, 1994). Nevertheless, the specific target sequence recognised by all Hox and HOM proteins *in vitro* is very similar (Hayashi and Scott, 1990). The most obvious explanation of these results is that the specificity of action of the Hox genes during embryo genesis is due to the presence of other co-factors (Manak and Scott, 1994). Vincenzo Zappavigna demonstrated in collaboration with our lab that HMG1 is indeed one of the co-factors that cooperates with the HOX proteins in the recognition of the specific target sequence. HMG1 and HOXD9 make protein-protein contacts via their HMG-boxes and homeodomain, respectively. By transient transfection assays, co-expression of HMG1 and HOXD9 enhances 4-fold the transcriptional activation of a reporter gene carrying the autoregolatory element recognised by HOXD9 (Zappavigna et al., 1996). In contrast, HMG1 is not able to increase the transcriptional activation mediated by another HOX protein, HOXD8, via the HOXD9-controlled element. Instead, HMG1 cooperates with HOXD8 when the cells are transfected with a reporter gene that contains the HOXD8-controlled element (Zappavigna et al., 1996). HMG1 is therefore a binding partner of all HOX proteins. HMG1 determines an increase of the binding affinity for DNA, and the transcriptional activation of the HOX proteins with the optimal target sequence.

1.5 HMG1/2 and site-specific recombination

As mentioned above, HMG1 and HMG2 are able to recognise cruciform DNA. This is a particular DNA structure similar to the letter X, that contains two angles of 60 degrees and two angles of 120 degrees. Generally, DNA assumes this conformation during recombination (Holliday junction), and during the extrusion of palindromic sequences under the effect of supercoiling (Lilley, 1990a; 1990b).

HMG1 and the prokaryotic histone-like protein HU have a similar function. HU is a small sequence-independent DNA binding protein involved in the transposition of the Mu phage. HU protein is capable of engineering DNA deformations required for the formation of higher order nucleoprotein structures (Lavoie and Chaconas, 1993; Lavoie and Chaconas, 1994). HU recognises with high affinity cruciform DNA (Pontiggia et al., 1993) and other DNA structures characterised by sharp bends or kinks, like bulged duplex DNAs containing unpaired bases. Moreover, HU strongly stimulates the site-specific recombination mediated by Hin recombinase of the prokaryote *Salmonella typhimurium* which mediates flagellar variation (Affolter et al., 1991; Lavoie and Chaconas, 1993; Lavoie and Chaconas, 1994). All these properties are analogous to the properties of HMG1: in fact HMG1 can substitute HU in the transposition of the Mu phage and in the site-specific recombination mediated by the Hin recombinase. The Integration Host Factor (IHF), is the prokaryotic correspondent of the HMG-box proteins that recognises specific sequences of DNA. IHF is a prokaryotic histone-like protein that is able to recognise a specific target sequence. As HMG1/2, IHF binds DNA in the minor groove and generates a bending in the DNA molecule (Drlica and Rouviere-Yaniv, 1987). It was demonstrated that the Leukaemia Enhancer Factor 1 (LEF-1), a HMG-box proteins of the sequence-specific class, can substitute IHF during the integration of the λ phage in the chromosome of *E. coli* (Giese et al., 1992).

1.6 HMG1/2 and V(D)J recombination

V(D)J recombination is a complex process that allows the generation of a nearly infinite number of antibodies and T Cell Receptors, which are the major molecular structures involved in antigen recognition. During V(D)J recombination, antigen receptor genes in B and T lymphocytes are assembled during development from V (variable), D (diversity), and J (joining) gene segments. The precise location of the initial cleavage reaction is directed by conserved recombination signal sequences (RSS) that flank coding segments of DNA. The RSS consists of two conserved sequence elements, a heptamer and an AT-rich nonamer, separated by a spacer whose length is either 12 or 23 bp. These two types of signals are referred to as 12- or 23-RSSs. Efficient recombination requires one of each type of RSS, a phenomenon known as the 12/23 rule (Lewis, 1994). The initial cleavage reaction of V(D)J recombination is catalysed by a combination of proteins encoded by the recombination activating genes (RAG1 and RAG2). The cleavage reaction consists of two steps. First, RAG1 and RAG2 cleave DNA at the RSS and determine the formation of free ends. Then, the free ends of coding sequences are linked to each other in a process that is carried out by factors that are involved in DNA repair (Schatz et al., 1989). RAG1 recognises the nonamer motif with low affinity but its interaction with RAG2 leads to the formation of a stable complex (RAG1/2 complex) that binds with specificity to the nonamer and heptamer sequences of the RSS (Difilippantonio et al., 1996; Spanopoulou et al., 1996). *In vitro*, the RAG1/2 complex binds with high affinity to the 12RSS while its interaction with the 23RSS is very inefficient and is assisted by HMG1 and HMG2 (Sawchuk et al., 1997; van Gent et al., 1997). The existence of a stable post cleavage complex with DNA containing signal ends was demonstrated: this complex not only contains RAG1 and RAG2, but also HMG1 or HMG2 (Agrawal and Schatz, 1997). *In vitro* recombination assays have demonstrated that the V(D)J cleavage is augmented in presence of HMG1 and HMG2 (Agrawal and Schatz, 1997).

The role of HMG1 and HMG2 in V(D)J recombination was recently investigated in our lab. The starting point of this work was the observation that the homeodomain of RAG1 (HD) showed a structural and functional similarity to the prokaryotic Hin recombinase DNA binding domain. Similarity also exists between the nonamer motif and the DNA target sequence of the Hin recombinase. Substitution of the RAG1 HD with that of the Hin-invertase produces a hybrid protein that is partially functional in V(D)J recombination (Spanopoulou et al., 1996). The functional parallels between RAG1 and the Hin recombinase on the one hand, and between HU and HMG1/2 on the other hand, suggested that the mechanisms by which HMG1 and HMG2 were involved in V(D)J recombination would be similar to that of HU and Hin recombinase in the site-specific recombination. We demonstrated that HMG1 and HMG2 physically interact with the RAG1 homeodomain (Aidinis et al., 1999). This interaction enhances the binding of the RAG1/2 complex to the nonamer motif. Moreover, we demonstrated that the recruitment of HMG1 and HMG2 facilitates the bending of the 23RSS by the RAG1/2 complex. The binding and bending of the 12RSS is not strongly facilitated by HMG1/2, probably because of the shorter distance between the heptamer and nonamer regions. The association of HMG proteins would stabilise the bent DNA, increasing the affinity of RAG proteins for the longer RSS. These observations suggest that HMG1 and 2 can serve as architectural elements in the assembly of higher-order nucleoprotein structures, as proposed for the prokaryotic HU and Hin recombinase. However, while in prokaryotes HU and the Hin recombinase exert their effect by direct binding to the DNA in the absence of any protein-protein interactions, in eukaryotes the HMG proteins have evolved the ability to both bend the DNA and to enhance DNA binding through direct protein-protein interaction.

1.7 HMG1 and HMG2 function

HMG1 and HMG2 interact physically with several homeodomain-containing transcription factors (HOX and OCT proteins), with steroid hormone receptors (SHRs), with components of the basal transcription machinery, with the V(D)J recombinase RAG1/RAG2 (Zwilling et al., 1995; Boonyaratanakornkit et al., 1998; Zappavigna et al., 1996; Agrawal et al., 1997; 1998). The interaction occurs between the HMG-boxes of HMG1 and 2 and the DNA binding domain of the partner, even in the absence of DNA. In the presence of DNA, the interaction between HMG1 and the partner facilitates the formation of the complex and/or enhances its stability.

The architectural capacity of HMG1 and HMG2 strongly suggest that these proteins are required for the formation or stabilisation of nucleoprotein complexes containing bent DNA, as shown in the model of Fig.5b. It has been hypothised that HMG1 or HMG2 can interact with a sequence specific protein prior to the interaction with DNA, and that the binding of the protein complex to the target DNA can induce a bending in the DNA molecule. Alternatively, either HMG1 or a sequence specific protein could first bind to the target DNA independently and facilitate the subsequent binding of the complementary module. In Fig.5 is also shown a DNA interaction model for HMG-box proteins of sequence-selective class, as SRY and LEF1. It has been postulated that these proteins recognise the target sequence and that induce a conformational change in the DNA molecule, favouring the contact of transcription factors distantly located on the DNA (Fig.5a).

1.8 The *hmg1* gene and the phenotype of HMG1-deficient mice

During my undergraduate work I participated to the cloning and characterisation of the mouse *hmg1* gene. The gene was isolated among a multitude of pseudogenes by virtue of the presence of introns. It is organised in five exons, the first is untranslated and the ATG is on the second exon. Exons 2 and 3 code for HMG-boxA, exons 3, 4, and 5

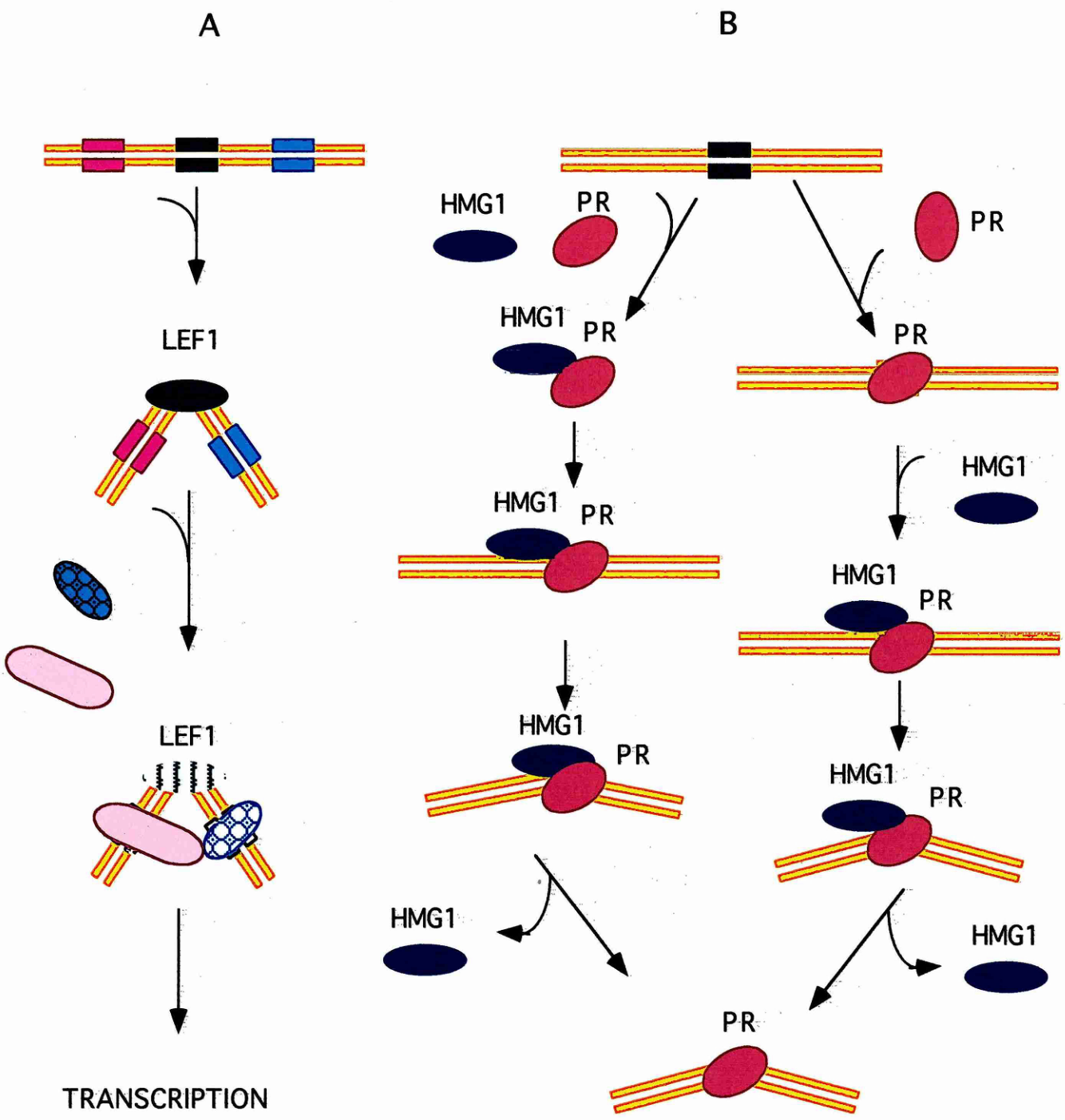


Fig.5a HMG-box of LEF1 binds a DNA target sequence and generates a bending in the DNA molecule that determines the interaction of transcription factors bound at distant positions. Fig.5b The binding of progesterone receptor (PR) on progesterone response element is efficient only in the presence of HMG1. In the absence of HMG1 PR don't bind the consensus sequence.

code for the HMG-boxB. Exon 5 codes for the acidic tail and contains a long 3'-untranslated region with three alternative polyadenylation sites (Ferrari et al., 1994). They determine the transcription of three different mRNA which are 2,4, 1,4 and 1 kb long; the most transcribed is the 1,4 kb one. HMG1 is ubiquitously expressed during development and in all adult tissues at a very high level.

hmg1 null mutant mice were produced in our lab, by conventional knock-out. The heterozygous mice were normal and produced offspring. *hmg1*$^{-/-}$ mice are born, but die within the first day of life. The immediate cause of death is a severe hypoglycaemia. The level of glucose is decreased in the blood of -/- animals, and intraperitoneal injection of glucose in -/- newborns permits the survival of some animals. However, the rescued animals do not recover completely to the normal phenotype, are very sick and anatomic analysis reveals the presence of atrophic organs and muscles, and the lack of fat tissue. Moreover, abundant glycogen was found in the liver of mutant mice, while it is absent in control mice because it is metabolised immediately after birth (Calogero et al., 1999).

Glucocorticoid-dependent gene expression is impaired in -/- mice. Fibroblasts from knock-out mice transiently transfected with a reporter plasmid containing the GR responsive element respond weakly to dexamethasone treatment, on the contrary in +/+ fibroblasts the reporter gene is activated. Moreover, $CD4^+/CD8^+$ T cells are partially resistant to dexamethasone treatment, while the exposure of +/+ thymocytes to dexamethasone determines massive apoptosis of double-positive $CD4^+8^+$ cells (Calogero et al., 1999).

In HMG1-deficient mice neither defects in the chromatin organisation nor defects in V(D)J recombination were found: the newborn mice had a normal serum abundance of immunoglobulins and thymus with a normal number of mature single-positive $CD4^+$ and $CD8^+$ T cells expressing a complete Vß repertoire (Calogero et al., 1999). Moreover, in spite of the cooperation between HMG1 and HOX proteins, the animals do not have any homeotic defects.

Surprisingly, HMG1 is not essential for the life of the cells, as it might be suggested by the abundance of the protein and its evolutionary conservation. Nevertheless, HMG1 is essential for the life of the mice after the birth, because it determines the proper transcriptional control of specific transcription factors (Calogero et al., 1999) (see Discussion).

The knock-out of HMG1 allowed new insights in the role of the protein and opens a new debate on the primary function of the HMG1/2 proteins. In fact, HMG1 seems to be necessary for its role as architectural regulator of transcription of certain genes, while appears to be not essential or vicariated in the other roles that were demonstrated *in vitro*.

1.9 HMG2: an overview

Biochemical and functional analysis reveal that HMG1 and HMG2 perform a very similar function. They have the same affinity for specific DNA structure, play the same role as co-activators of different transcription factors, and are both involved in V(D)J recombination.

The human *HMG2* gene was isolated two years before the mouse *hmg1* gene. It has the exon-intron organisation typical of HMG1/2 family genes, as described above for mouse *hmg1* (Shirakawa and Yoshida, 1992). Discordant hypotheses were formulated on the possible function of HMG2. Some authors noted that the level of HMG2 parallels the proliferative activity of organs such as thymus, spleen, and testis, and suggested that HMG2 plays a role in cell replication (Seyedin and Kistler, 1979). Other evidence is in agreement with this hypothesis: the level of the HMG2 mRNA is higher in exponentially growing cells and in cells transformed with various viral genes (Shirakawa and Yoshida, 1995; Yamazaki et al., 1995). Moreover, cell growth is repressed by expression of antisense HMG2 RNA (Yamazaki et al., 1995). These findings suggest that the HMG2 protein might be linked to cell proliferation. In contrast, other authors analysed the distribution of HMG2 in different adult tissues from different species and observed three

levels of HMG2 content: low, intermediate and high (Bucci et al., 1984; 1985). All non-proliferating fully differentiated tissues had low contents of HMG2. All proliferating tissues or those that possessed a high proportion of lymphoid tissue (capable of rapid proliferation after stimulation) had intermediate contents of HMG2. The testis was the only rat tissue with a high content of HMG2. Moreover, the authors analysed the distribution of HMG2 during spermatogenesis in rat (Bucci et al., 1984; 1985), and found that the increase of HMG2 levels is associated with pachytene spermatocytes and early spermatids, while levels in spermatogonia and early spermatocytes were low as in the somatic cells of testis (Leydig and Sertoli cells). They discovered that the increment of HMG2 level in testis is due to a faster migrating form of HMG2. Since pachytene spermatocytes are non-proliferating cells that undergo genetic recombination and very active RNA synthesis (Bellve et al., 1977), they concluded that HMG2 is not associated with cell proliferation (Bucci et al., 1984; 1985).

The literature on HMG2 function is rather confuse, and the analysis of the protein distribution is also contradictory (Bucci et al., 1984; 1985; Seyedin and Kistler, 1979; Shirakawa and Yoshida, 1995). For this reason I decided to clone and characterise the mouse *hmg2* gene and to inactivate it *in vivo*. The purposes of my work were also to analyse *hmg2* expression in the mouse, both during development and in adult tissues, using techniques different from those used in previous works, and to describe possible differences between HMG2 and HMG1.

2 RESULTS

2.1 Isolation of the mouse *hmg2* gene

We isolated the mouse *hmg2* gene by means of a screening procedure based on the amplification of introns by PCR, that we previously used for the cloning of the mouse *hmg1* gene in order to distinguish it from a multitude of processed pseudogenes (Ferrari et al., 1994). We adopted this strategy although the presence of hmg2-related pseudogenes was not demonstrated. We assumed that the structure of the mouse *hmg2* gene would be similar to that of the human *HMG2* gene, which contains four introns (Shirakawa and Yoshida, 1992). We synthesised three pairs of oligonucleotides corresponding to the HMG2 cDNA sequence flanking putative introns 2, 3 and 4, and ending at their 3' terminus with one or two bases corresponding to the canonical splicing junction 5'-GT and 3'-CT. We obtained the amplification of the introns 3 and 4 (INT3 and INT4), which are 120 and 540 base pair long, respectively. A PCR reaction performed on mouse genomic DNA using the primers 5' to intron 3 and 3' to intron 4 gave the amplification of a 800 bp fragment, indicating that INT3 and INT4 are collinear on the mouse genome. In Fig.6 the experimental strategy we used to amplify intron 4 is shown.

Using the INT4 as a probe we screened a 129Sv mouse genomic library and we obtained 20 positive clones from 10^6 phage plaques. 16 of these clones were positive to the amplification of INT3 and presented overlapping restriction maps. *Eco*RI fragments from two of the positive clones were subcloned into the pBlueScriptII KS(+) plasmid (pBS-KS+) and characterised.

The *hmg2* gene was completely sequenced: small sized fragments of the gene were subcloned in the pBS-KS+ vector and were sequenced using the canonical primers T3 and T7. The sequence includes the 5'-untranslated region, the coding region of the mouse HMG2 cDNA and a long 3'-untranslated region (UTR) (Fig.7).

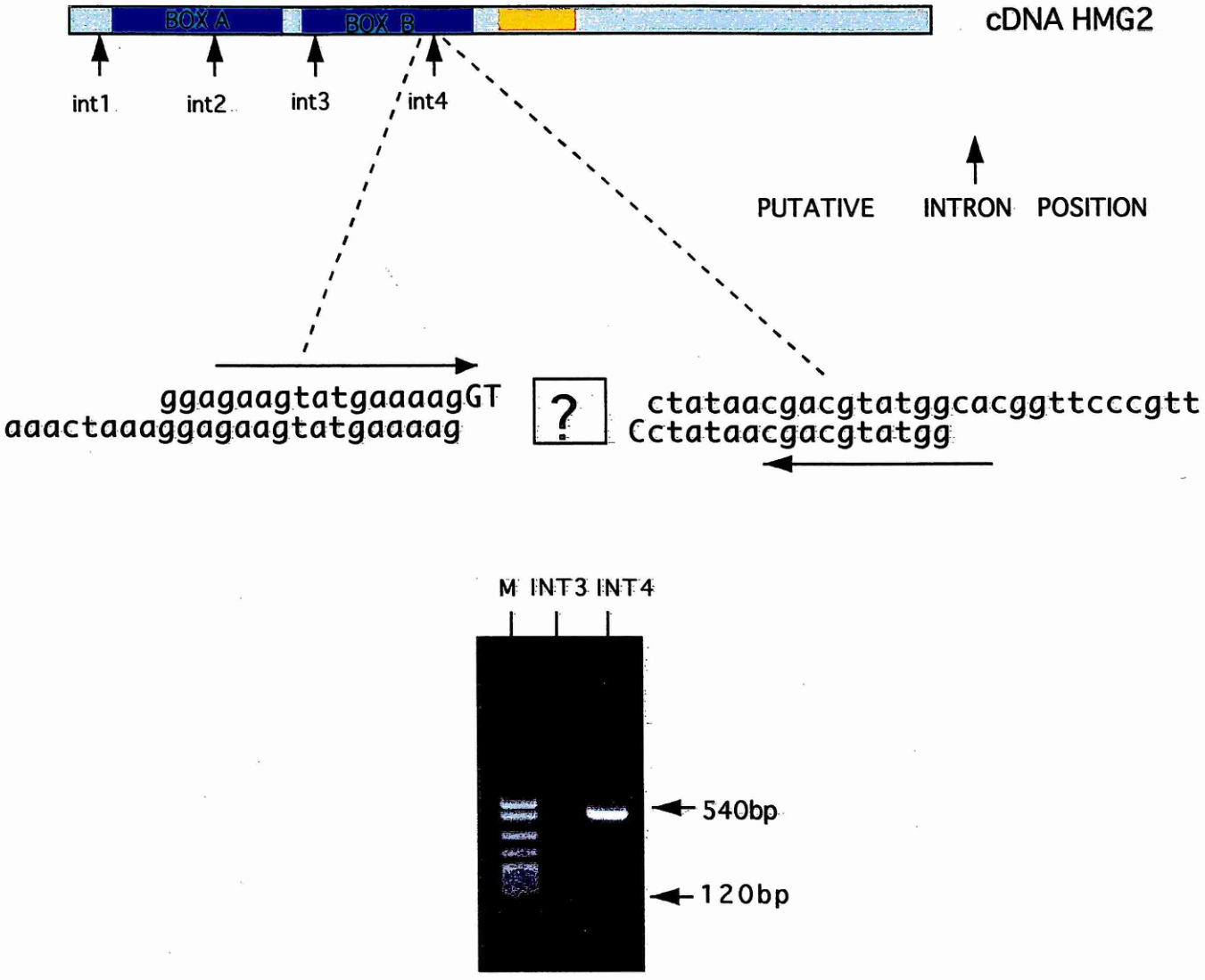


Fig.6 Experimental strategy used for the amplification of intron 4; the sequence of the oligonucleotides are indicated; they contain one or two additional bases at their 3' terminus corresponding to the canonical splicing junction.
PCR performed on mouse genomic DNA using the primers described above gave the amplification of a fragment 540 bp long (arrow). Using the primers flanking the putative intron 3 we obtained the amplification of a 120 bp fragment (arrow).

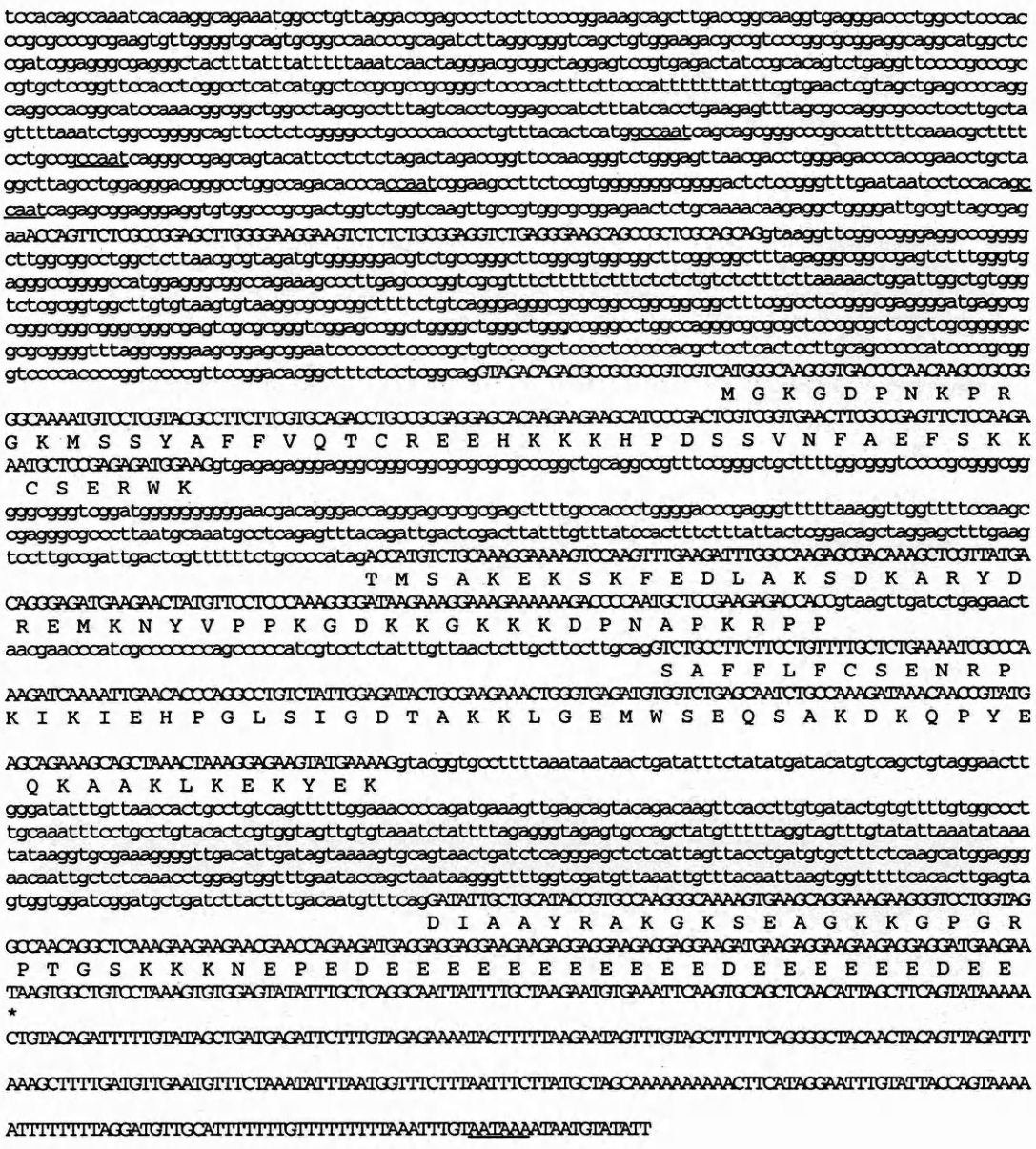


Fig.7 Sequence of the *hmg2* gene. Exon sequences are shown in UPPERCASE. The CCAAT boxes are underlined. The deduced amino acid sequence is indicated below the coding region, and an asterisk indicates the stop codon. The putative polyadenylation signal, AATAAA, is underlined.

There is no difference between the coding region of the mouse *hmg2* gene we isolated and the sequence of the cDNA described by Zwilling et al.(1995). The four introns were also completely sequenced; they are fairly typical of introns, because they are rich in AT, lack any open reading frame and have consensus splicing sites. The introns at positions 73-644, 819-1133, 1280-1361, and 1537-2037 are 571, 314, 81, and 500 bp long, respectively. The 3'-UTR contains a putative polyadenylation signal, AATAAA at positions 2542-2547, that falls in a region rich in AT that is 18-13 bp upstream of the 3'-terminus of exon 5.

2.2 Structure of the mouse *hmg2* gene

The mouse *hmg2* gene is composed of 5 exons, the first of which is untranslated and falls in a region of very high C and G content with the features of a CpG island. Exon 2 carries the translation start site. The HMG-boxA is encoded by exon 2 and exon 3, and the HMG-boxB is encoded by exon 3, 4 and 5. Exon 5 also codes for the terminal acidic tail, and contains a long 3' untranslated region. Fig.8a shows the structure of the gene and the previously described structure of the genes that compose the HMG1/2 family: the mouse *hmg1* and the recently found mouse *hmg4* gene (Ferrari et al., 1994; Vaccari et al., 1998). These genes have a very similar structure and the position of introns is absolutely conserved, suggesting that the genes arose from an ancestral duplication (Fig.8b) (see the Introduction).

The transcription start site was identified by RNase protection assay performed on total RNA from embryonic fibroblasts. We obtained the protection of two fragments, which are 72 and 77 bp long, respectively (Fig.9). The shorter fragment coincides with the 5' terminus of the HMG2 cDNA, the longer is extended 5 base pair upstream and represents an alternative transcription start site. The following reasons could account for the relative low intensity of the protected HMG2 fragment as compared to the ß-actin control: the expression level of the *hmg2* gene in this cell type may be low and the

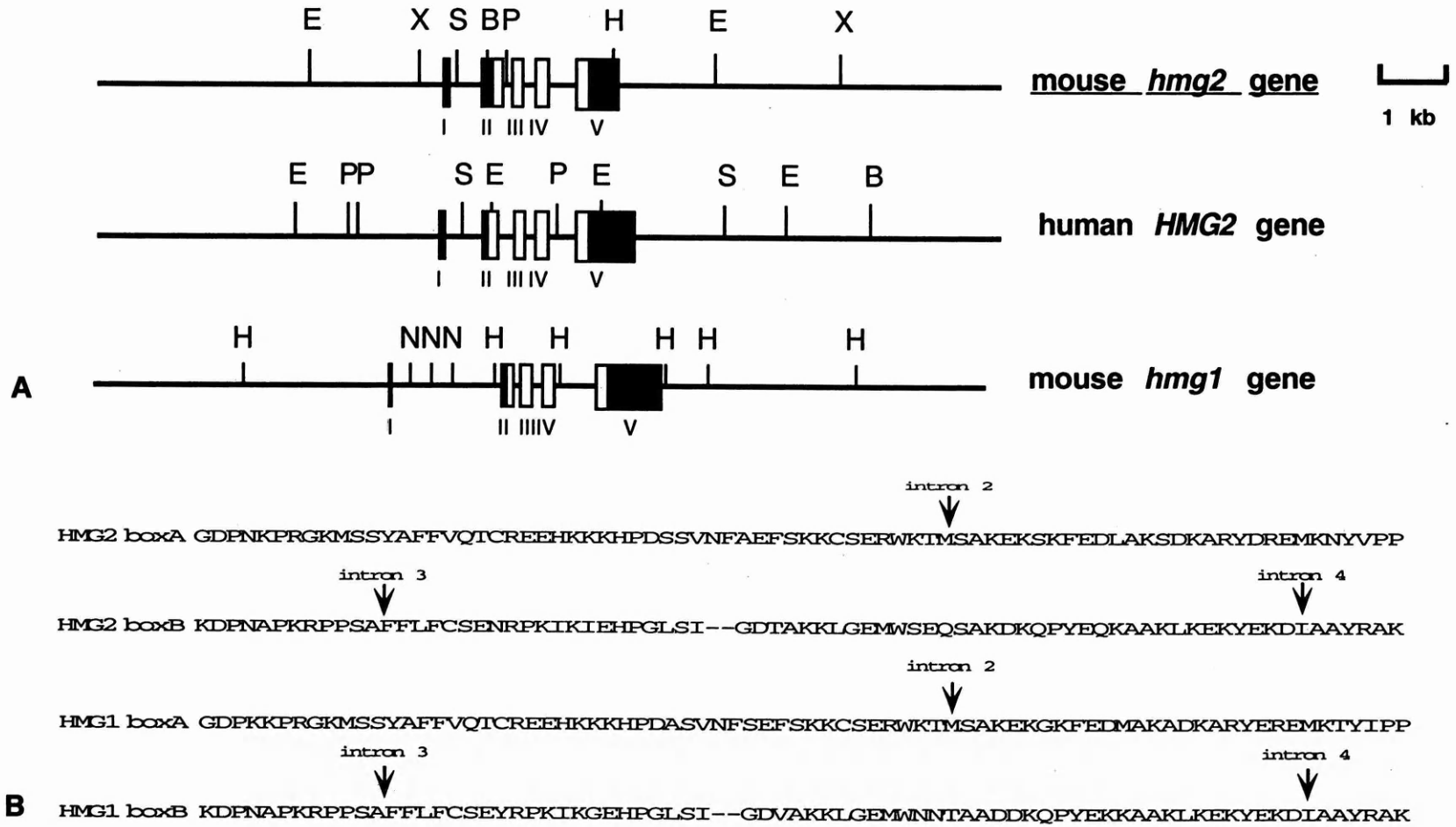
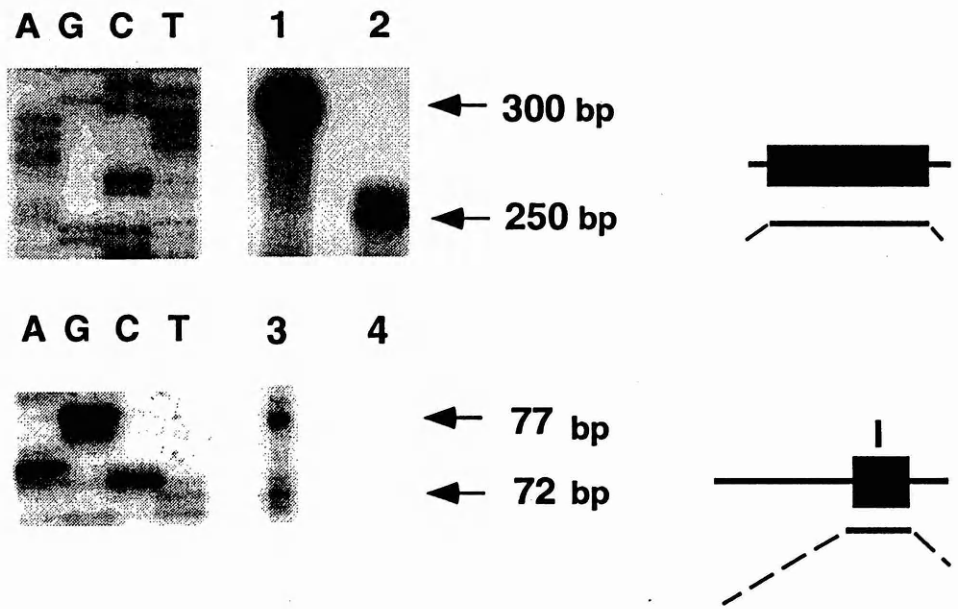


Fig.8 (a) Structure of the *hmg2* gene. Exons are indicated as boxes (open for the coding region and filled for the untranslated region). The structure of the gene was compared to the human gene and the mouse *hmg1* gene. The genes of the HMG1/2 family are highly conserved. (b) sequence alignment of HMG-boxA and B. The locations of introns are indicated by arrows. The positions of introns is conserved between mouse HMG1 and 2, and between human and mouse HMG2 (not shown).



caatcagagcggagggaggtgtggcccgcgactggtctggtcaagttgccgtggcgcggagaactctgcaaaacaagaggctggggattgcgttagcgag

aaACCAGTTCTCGCCGGAGCTTGGGGAAGGAAGTCTCTCTGCGGAGGTCTGAGGGAAGCAGCCGCTCGCAGCAGgtaaggttcggccgggaggcccgggct

tggcggcctggctcttaacgcgtagatgtgggggacgtctgccgggcttcggcgtggcggcttcggcggctttagagggcggccg

Fig.9. The transcription start site of the *hmg2* gene was determined by RNase protection on total RNA from embryonic fibroblasts. The riboprobe was obtained by in vitro transcription of a 270-bp SmaI-XbaI fragment of the *hmg2* gene, around exon1, and cloned in pBS-KS+. Two protected fragments of 72 and 77 bp were obtained using the antisense probe (lane3), while no protected fragment was obtained using the sense probe (4). Lane 1 and 2 correspond to the β-actin protected and unprotected fragments, respectively. At the bottom we can see the sequence of exon 1 (uppercase), the spots correspond to the transcription start sites.

protected fragment is shorter than the ß-actin fragment and contains a high percentage of GC. Moreover, the riboprobe we used protects a region that is more extended upstream than downstream of the exon 1, so there is a long region of the riboprobe that is not matched to the mRNA and therefore it is more subject to digestion by the ribonuclease (see Materials and Methods).

Transcription terminates at a single site: in fact a single hmg2 transcript was found by Northern blot analysis in all mouse tissues and cell lines (Fig.16a) and a single polyadenylation site has been detected in cDNA (Zwilling et al., 1995).

2.3 The promoter of the mouse *hmg2* gene

The sequence upstream of the transcription initiation sites contains no TATA box, while CCAAT sequences are present at positions -340, -295, -165, and -103. To test if the promoter region is able to drive transcription, I constructed a plasmid, GEPPETTO, in which a fragment encompassing 1.2 kb upstream of the transcription start site, exon 1, intron 1 and part of exon 2, was fused in frame to the *lacZ* gene. The expression of this construct results in the production of a chimeric protein capable of cleaving the ONPG substrate (see Materials and Methods). The plasmid was introduced in NIH3T3 cells and ßGal activity was measured 48 hours after transfection. Fig.10 shows that GEPPETTO is able to induce transcription of the *lacZ* gene at a level comparable to the positive control, pUT529Δ, that contains the selectable gene under the control of the TK promoter and the Polyoma enhancer. Sequence comparison between human and mouse *hmg2* gene reveals high homology from nucleotide -1 to nucleotide -425 (Fig.11), suggesting that this region contains all the elements necessary for the activation of the gene. For this reason I derived from GEPPETTO two constructs with progressive deletions, GEPPO that contains the whole homologous region and GEPPINO that contains a region 175 bp shorter than GEPPO. As shown in Fig.10, GEPPINO induces transcription at low level, as

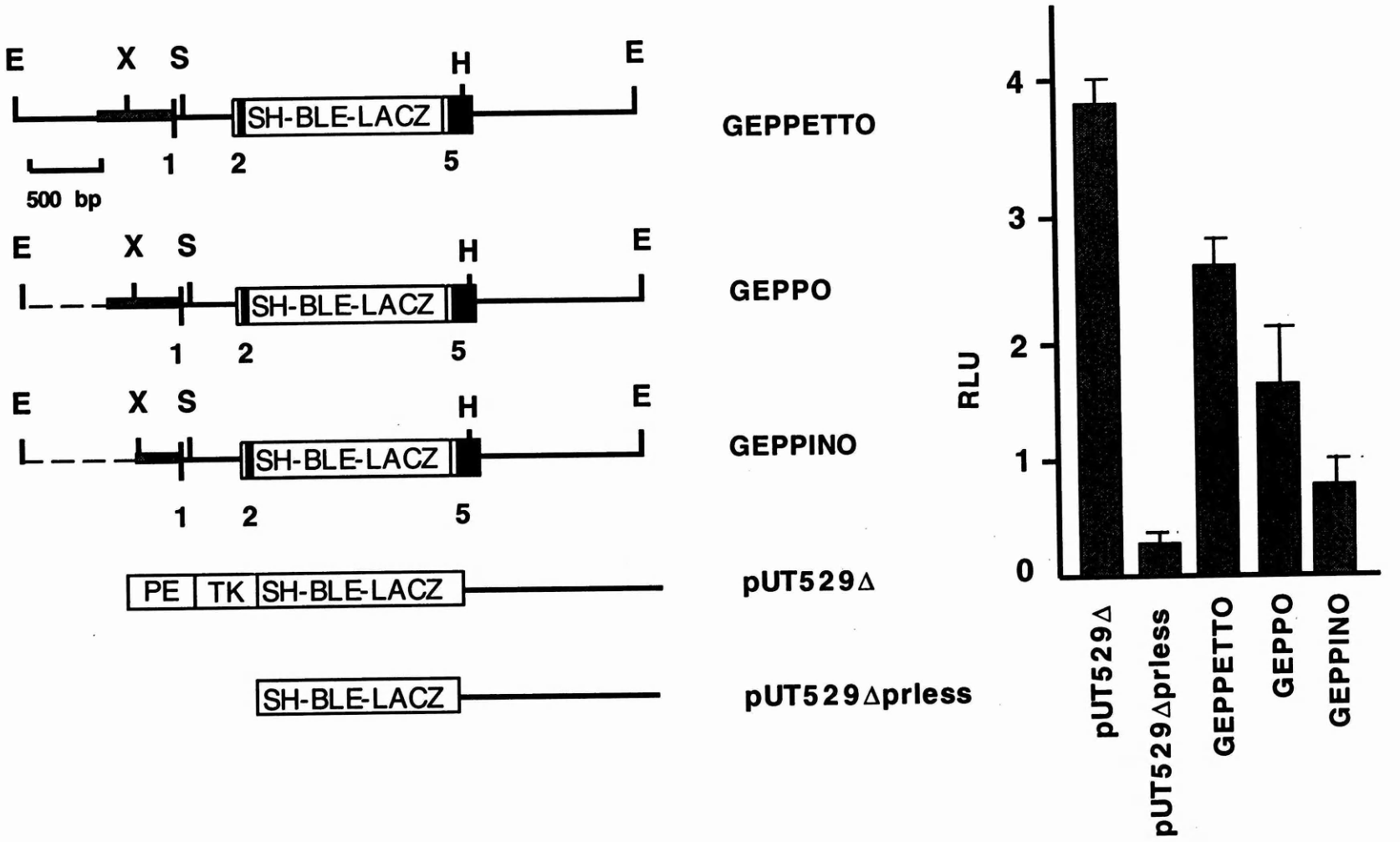


Fig.10 Left: constructs carrying different deletions of the promoter region of *hmg2*. GEPPO is the construct that carries the region of homology with the promoter of the human *HMG2*, GEPPINO and GEPPETTO are two constructs that carry respectively a shorter and longer promoter region in comparison to GEPPO. Right: NIH3T3 cells were transiently transfected with 6 μg of each construct and β-gal activity was measured. Contrary to expections, GEPPETTO drives transcription better than GEPPO. pUT529Δ and pUT529Δ promoter less were used as positive and negative controls. PE, polyoma enhancer. TK, TK promoter. The grey box represents the region of homology between human *HMG2* and mouse *hmg2* promoters. E, *Eco*RI; X, *Xba*I; S, *Sma*I; H, *Hind*III.

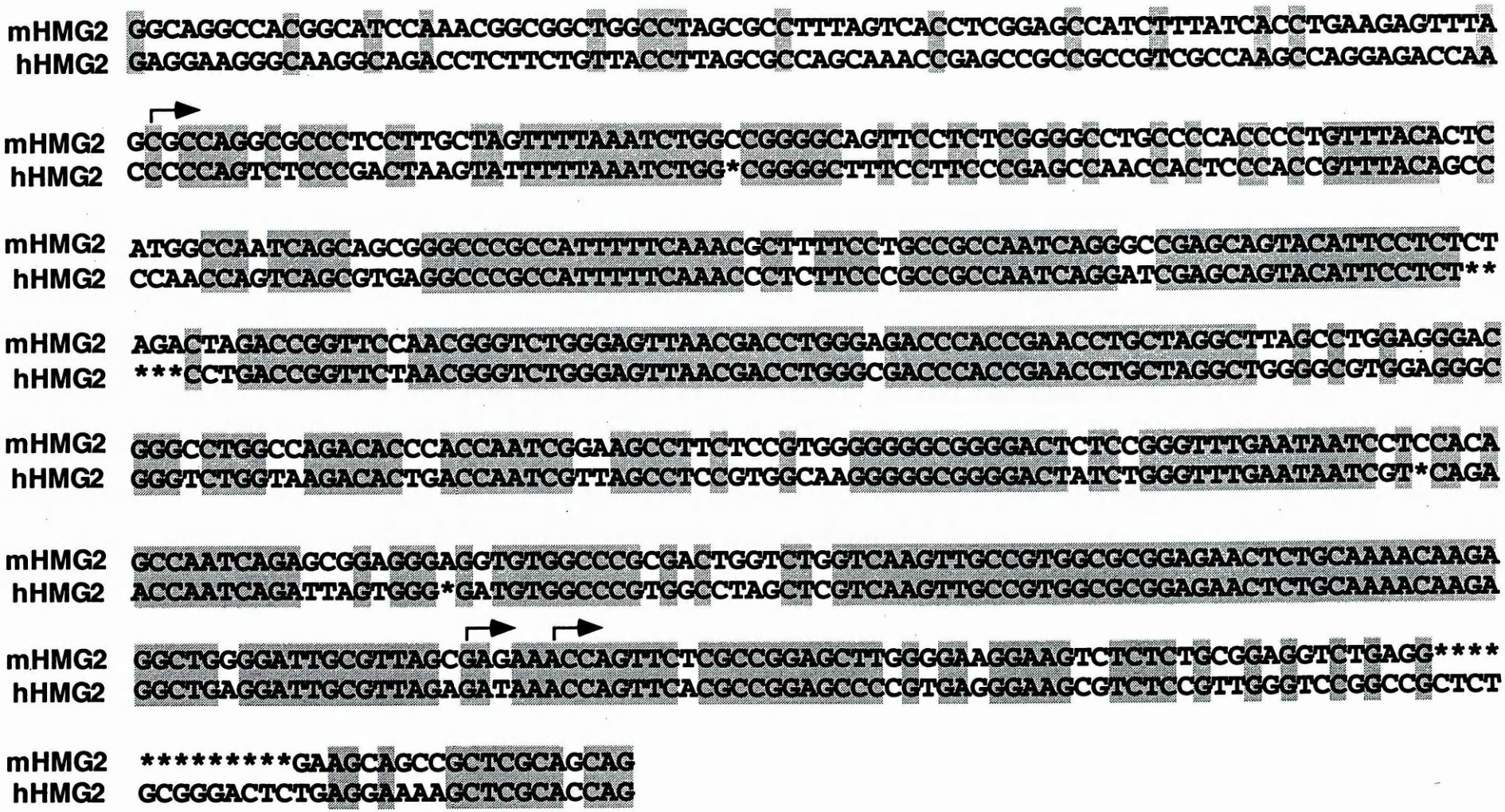


Fig.11. Sequence comparison between the promoter region of the mouse and human HMG2 gene. The homology is very high and probably this region contains all the promoter elements of the two genes. The first arrow indicates the start of homology, the other two arrows indicate the transcription start site of the mouse *hmg2* gene.

expected since it contains a short region of the promoter. Unexpectedly, GEPPO is not able to induce transcription at same level as GEPPETTO, indicating that the region upstream probably contains additional elements involved in the control of the HMG2 promoter.

2.4 Mapping of the mouse *hmg2* gene

To map the mouse *hmg2* gene we looked for a polymorphism between recombinant inbred strains. A polymorphism can be due to the presence or absence of particular restriction sites, to differences in the sequence and to the presence or absence of specific microsatellites. Introns and 5'- and 3'-untranslated regions are generally not conserved between different strains, while coding regions are generally conserved. We decided to look for polymorphisms in the introns. We amplified by PCR intron 4 on genomic DNA from C57BL/6J and *M.spretus*, respectively. PCR products were run on a non-denaturing polyacrylamide gel, and we found a polymorphism on the length of intron 4. We then mapped the mouse *hmg2* gene by linkage analysis of the BSS backcross panel distributed by The Jackson Laboratory (see Materials and Methods). A radioactive PCR was performed on two parentals and 94 N_2 DNA obtained from the backcross progeny of (C57BL/6J X *M. spretus*) F1 X *M. spretus*. Fig.12a shows a representative polymorphic pattern we obtained. The ensuing Strain Distribution Patterns (SDP) were analysed with Map Manager 2.6.5. (Mainly and Elliott, 1991), leading to map localisation and linkage analysis with respect to markers independently mapped by others. These analysis allowed us to unequivocally assign the *hmg2* to the centromeric portion of chromosome 8 (Fig.12b and c). These region is syntenic to the human chromosome band 4q31 where the human HMG2 maps (our unpublished results and (Wanschura et al., 1996).

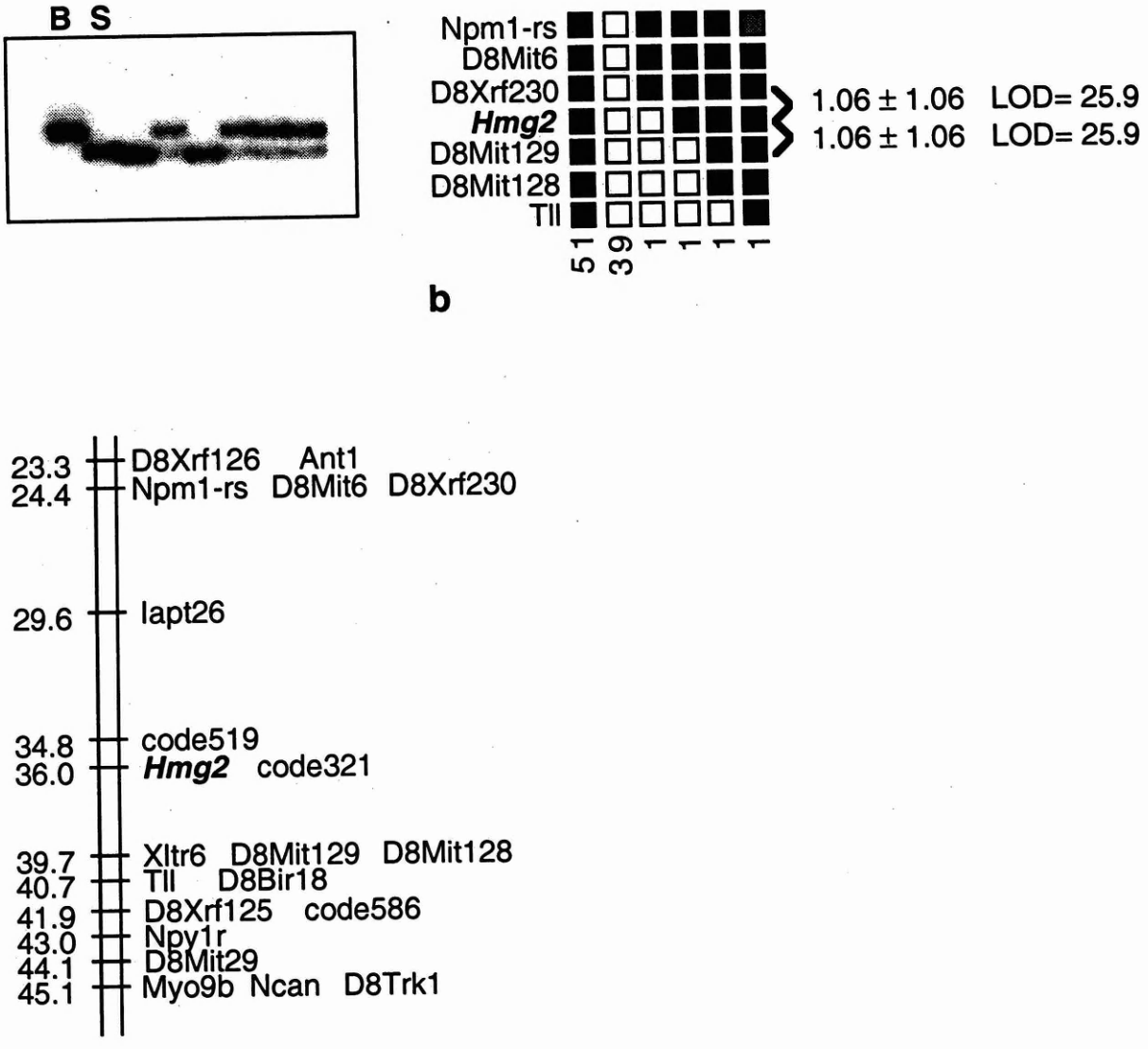


Fig.12 (a) Representative polymorphic banding pattern obtained amplifying intron 4 on DNA from BSS backcross. B, B6 parental; S, *Mus spretus* parental. (b) Genetic mapping of *hmg2* on mouse chromosome 8. Rows represent strain distribution patterns. Empty squares indicate the *Mus spretus* allele; filled squares indicate B6 alleles; stippled squares indicate genotype not determined. Number between rows indicate recombination fractions +/- standard error. Columns represent different haplotypes observed on chromosome 8. Numbers below columns define the number of individuals sharing each haplotype. (c) Position of *hmg2* on chromosome 8 with respect to markers independently mapped by others on the BSS backcross.

2.5 HMG2 enhances the transcriptional activation mediated by HOXD9

After the cloning and the characterisation of the mouse *hmg2* gene I did some functional experiments to better understand the role of the protein in the cell. The experiments were chosen to point out possible similarities or differences between HMG1 and HMG2 function on one hand, and to facilitate the phenotypic analysis of the HMG1 deficient mice, on the other.

It was previously shown, in a work where I collaborated, that HMG1 and HMG2 act similarly as architectural regulators of the transcription mediated by steroid hormone receptors (SHRs) (Boonyaratanakornkit et al., 1998). Both HMG1 and HMG2 physically interact with estrogen, androgen, glucocorticoid and progesterone receptors, while interaction with nonsteroid nuclear receptors was not detected. The protein-protein interaction occurs between HMG-boxes and DNA binding domain (DBD) of SHRs, and significantly increases binding affinity for DNA of HSRs. The increase of binding affinity of HSRs for DNA induced by HMG proteins *in vivo* corresponds to an enhancement of 7 to 10 fold of HSRs-mediated transcriptional activation (Boonyaratanakornkit et al., 1998). No differences between HMG1 and HMG2 were found, both in the protein-protein interactions and in the *in vivo* assays, suggesting that both the proteins play the same role in this specific process.

A role similar to that described above was suggested for HMG1 in another context. Vincenzo Zappavigna demonstrated in collaboration with our lab that HMG1 physically interacts with HOXD9, and enhances the affinity of HOXD9 for DNA and HOXD9-mediated transcriptional activation (Zappavigna et al., 1996). In our lab, Tiziana Bonaldi demonstrated that HMG2 as well establishes protein-protein contacts with HOXD9. For this reason, I decided to analyse the functional role of these interactions in cell transfection assays, in order to verify whether also in this context HMG1 and HMG2 have the same behaviour. The *HOXD9* gene product activates transcription in transient

co-transfection assays in NIH3T3 through an autoregulatory element (HCR) that is evolutionary conserved (Zappavigna et al., 1991). This element constitutes the binding site for the HOXD9 protein. We used a luciferase reporter, pTHCR, where the HCR element is fused to the herpes simplex virus thymidine kinase (TK) promoter. To test whether the pHMG2 expression construct, that contains the whole gene under the control of its own promoter, is able to stimulate reporter activity, we transiently co-transfected NIH3T3 with constant amounts of pHMG2 and pTHCR and increasing amounts of the HOXD9 expression construct, pSGD9 (See Materials and Methods). pHMG2 was able to stimulate the transcription of the reporter to a level 6 to 7-fold higher than the maximum level obtained by transfection with pTHCR and pSGD9 alone (Fig.13a). Fig.13b shows the enhancement of the activity induced by 1.5 or 5 μg of pHMG2 when co-transfected with increasing amounts (0.6-5 μg) of pSGD9. The transactivation is enhanced linearly with the amount of transfected HOXD9 and the enhancement is absent in transfection with pHMG2 alone. Fig.13 also shows the comparison between the activity induced by HMG2 and the one induced by HMG1. 5 μg of HMG2 and HMG1, respectively, were co-transfected with increasing amounts of HOXD9. HMG2 seems to be a more potent co-activator than HMG1; the enhancement of the activity mediated by HMG2 is 2.5-fold higher than that mediated by HMG1. In order to understand the reason for the different behaviour between HMG1 and HMG2 I compared the amino acid sequence of the two proteins. The major difference I found was the length of the acidic tail: HMG2's is 25 aa long while HMG1's is 30 aa long. In our lab we found that truncated forms of HMG1 and HMG2 lacking the whole acidic tail interact with other proteins (including HOXD9) better than the wild type forms. This data suggests that the acidic tail plays a negative role in protein interaction and probably also in transcriptional activation. To test this hypothesis I derived two construct from pHMG2: one carries the deletion of the acidic tail (pHMG2Δ) and the other one carries the acidic tail of HMG1 instead of that of HMG2 (pHMG2switch, see Materials and Methods and Fig.14). Fig.14 shows the result

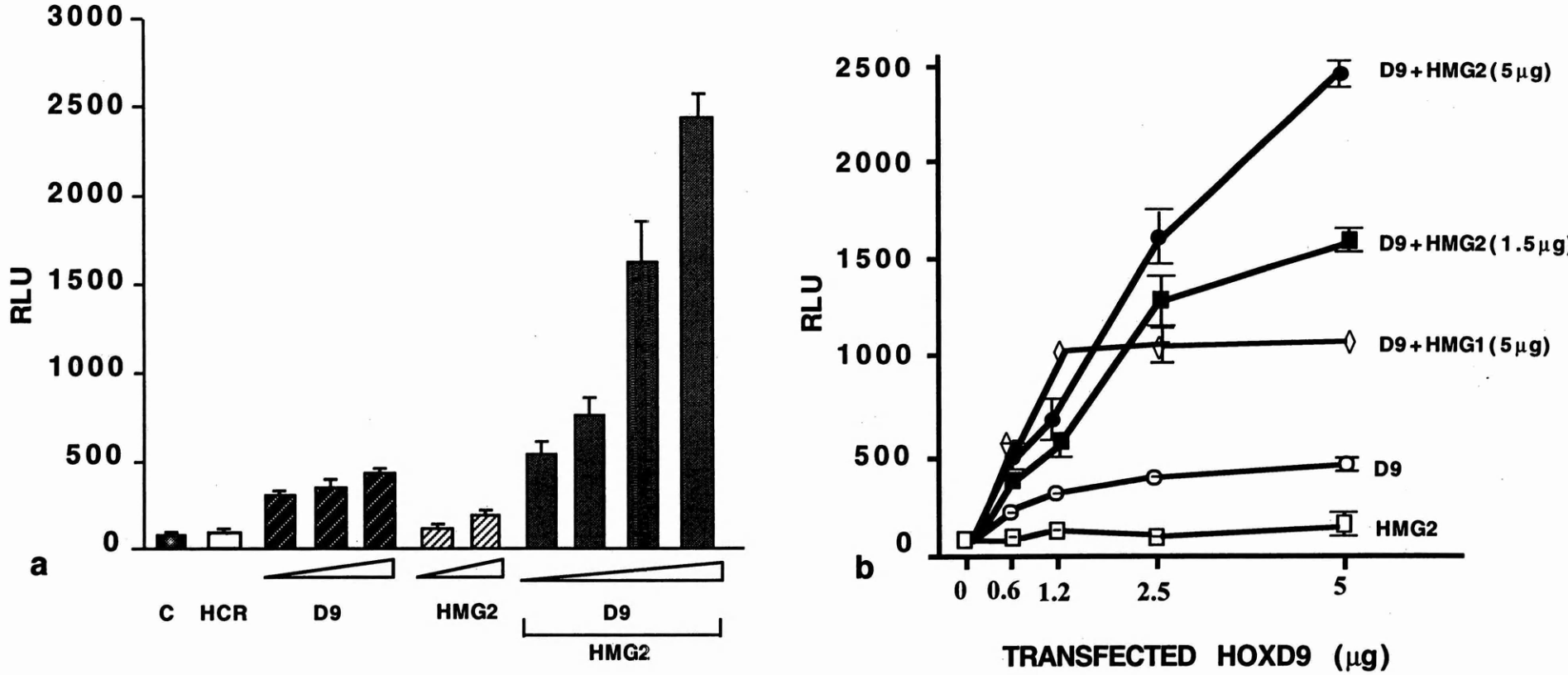


Fig.13 (a) HMG2 enhances the transcriptional activity of HOXD9 in transient co-transfection assays. NIH3T3 cells were transfected with 6 μg of pT81luc (C) or the pTHCR (HCR) luciferase reporter constructs, together with increasing amounts (0.6, 1.5, 3, and 5μg) of the HOXD9 (D9) expressor construct, or with increasing amounts of pHMG2 (0.6, 1.5, 3, and 5μg). (b). Effect of HMG2 in presence of various amounts of HOXD9. Various amounts of HMG2 alone (open squares), of HOXD9 alone (open circle), or of HOXD9 with 1,5 or 5 μg of pHMG2 (filled squares and filled circle), or of HOXD9 with 5μg of HMG1 (open diamonds) were transfected together with the pTHCR reporter. RLU, relative luciferase units. Cells were cotransfected with pRSVβgal and luciferase values were normalised with βgal activity.

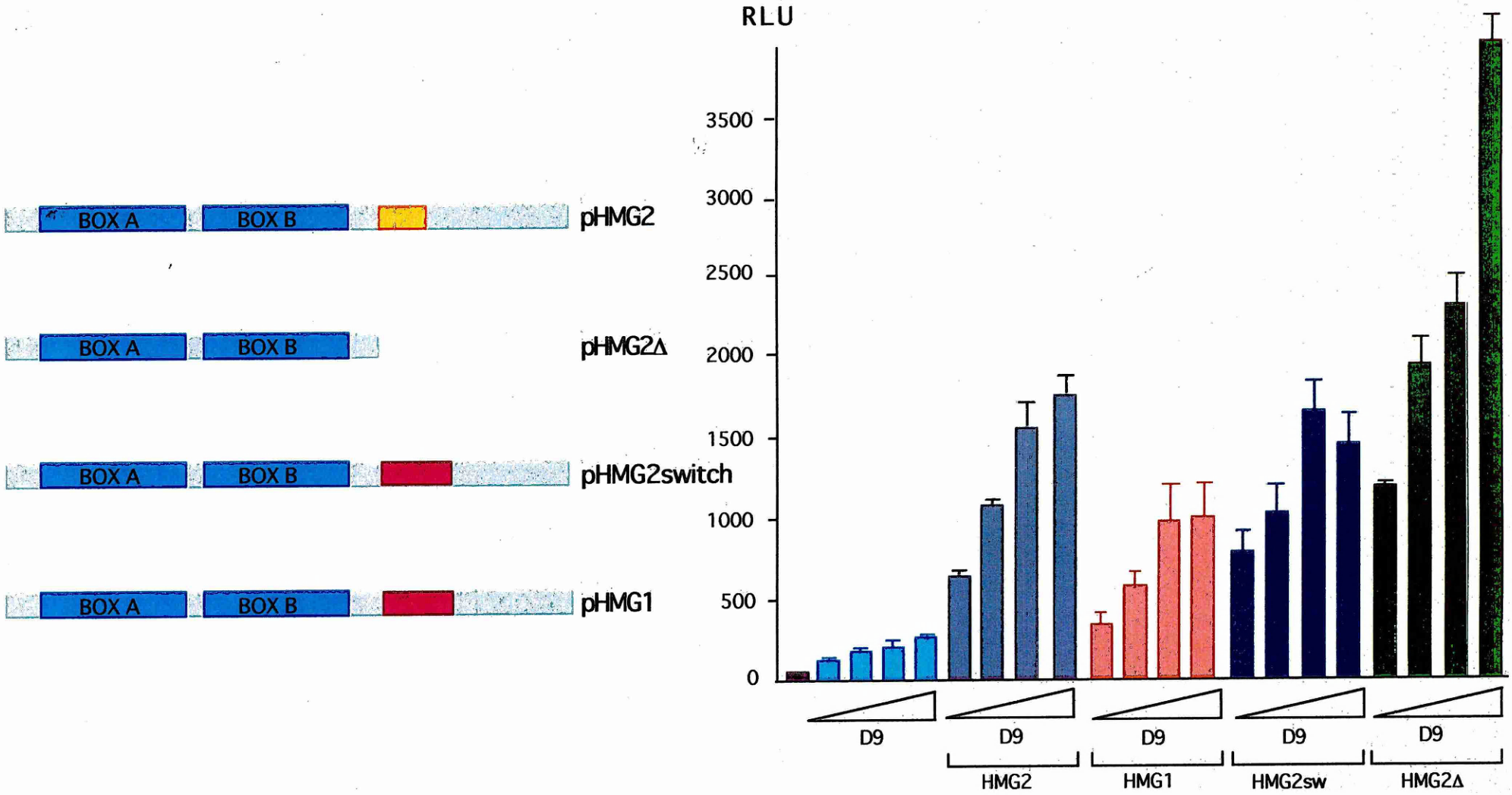


Fig.14. Role of the acidic tail of HMG2 in the transcriptional activation mediated by HOXD9. Left: different costructs produced from pHMG2. pHMG2Δ carries the deletions of the acididc tail, while pHMG2switch carries the acidic tail of HMG1. Right: NIH3T3 were transfected with increasing amounts of HOXD9 (D9) (0.6, 1.5, 3, and 5 μg) and 6 μg of pTHCR (HCR) reporter plasmid, and 5 μg of expressor constuct (pHMG2, pHMG1, pHMG2switch, or pHMG2Δ)

obtained: pHMG2swhitch contrary to the hypothesis works like HMG2 wild type, indicating that probably differences in the acidic tail cause no functional differences between HMG1 and HMG2. On the contrary, when we transfected cells with pHMG2Δ we obtained a significant increase in the transcriptional activation: 2.5-fold more than the activation in cells transfected with the wild type protein and around 15-fold more than those in cells transfected with pTHCR and pSGD9 alone (Fig.14). The data presented above show that the acidic tail does play a role in the transcriptional activation mediated by HOX proteins.

2.6 HMG2 is not stably associated to chromosomes

In our lab we showed that the association of HMG1 with chromatin is markedly different from that of histone H1 (Falciola et al., 1997; Spada et al., 1998). These results were in contrast to the prevailing view that HMG1 is a structural component of chromatin and to the observations that during *Drosophila* and *Xenopus* embryogenesis HMG1 appears to substitute for histone H1 (Dimitrov et al., 1993; 1994; Ner and Travers, 1994) (see Introduction). I analysed the association of HMG2 with chromatin, in order to verify whether its behaviour is similar or different to that defined for HMG1. I used the same conditions described in Falciola et al. (1997). The association of HMG2 with chromatin was analysed in both interphase and metaphase fibroblasts. NIH3T3 cells were treated with nocodazole, an inhibitor of microtubule assembly: mitotic cells were shaken off the dishes, while non-mitotic cells were detached by mild trypsin digestion. The two cell populations were then permeabilised with NP-40 and analysed for protein retention by Western blotting. HMG2 is detached from mitotic chromosomes, contrary to histone H1 and HMG-I(Y), another chromatin component. During interphase, HMG2 diffuses out of nuclei after permeabilisation of the nuclear membrane, whereas histone H1 and HMG-I(Y) remain associated to chromatin (Fig.15).

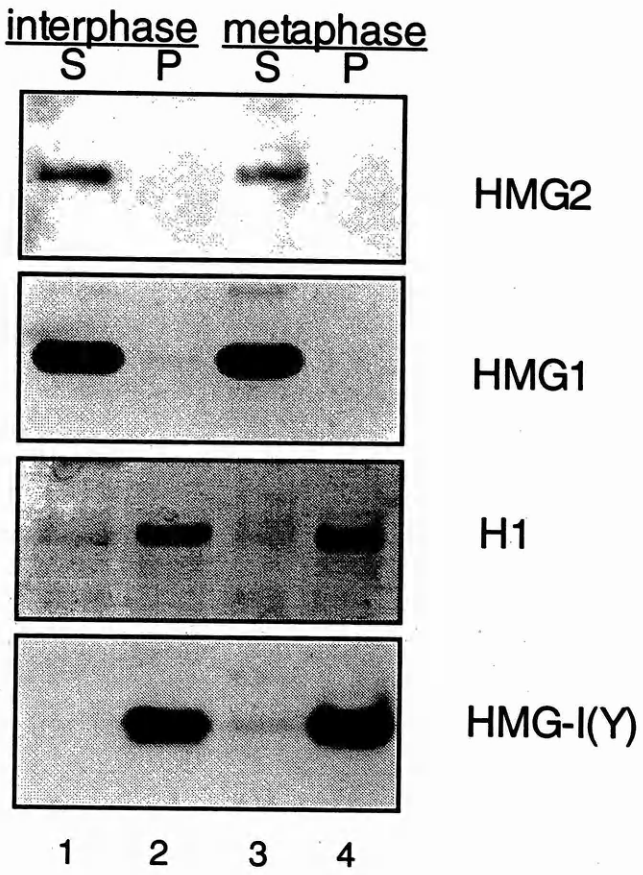


Fig.15 HMG2 is not stably associated with chromatin and condensed chromosomes. NIH3T3 fibroblasts were exposed overnight to nocodazole, a drug that inhibits microtubule polymerisation. Mitotic cells were shaken off the dishes. Both metaphase and interphase cell populations were permeabilised with 0.1% NP-40 and immediately centrifuged. Supernatants (lanes 1 and 3) and cell pellets (lanes 2 and 4) were analysed by Western blotting. Both in interphase than in metaphase cells HMG2 and HMG1 are released into the medium. Extracts were also analysed for two chromatin proteins: histone H1 and HMG-I(Y). In both mitotic and nonmitotic fibroblasts histone H1 and HMG-I(Y) are retained in the cell pellets (lanes 2 and 4).

HMG2 protein is not stably associated with chromatin, neither in interphase nor in metaphase cells. Therefore, HMG1 and HMG2 interact with chromatin in the same way.

2.7 Expression pattern of the mouse *hmg2* gene

I analysed the distribution of the protein in several adult mouse tissues by Western blot. Surprisingly, HMG2 is not detectable in liver, brain, muscle, and uterus (Fig.16a). Only testis contains high levels of HMG2. Kidney and spleen contain intermediate levels of protein.

I further analysed the expression of the mouse *hmg2* gene by Northern analysis and by *in situ* hybridisation of pre- and postnatal tissue sections.

By RT-PCR I detected *hmg2* in embryonic stem cells, that derive from a blastocyst 4 days *post coitum* (not shown). The expression of *hmg2* probably begins early, since zygotic transcription of *hmg1* was shown to begin at the 2-cell stage (Spada et al., 1998).

We performed Northern analysis using the HMG2 cDNA as probe. A single mRNA species of approximately 1 kb was detected in embryonic and some adult mouse tissues. Fig.16b shows that the *hmg2* gene is highly expressed in the three different stages of development analysed, embryonic day (E) 10.5, 14.5 and in liver, brain and carcass of E18.5 embryos. hmg2 transcript levels strongly decline in the adult tissues. In particular, hmg2 messenger RNA is not detectable in adult liver and brain, while it is highly expressed in adult thymus, spleen and testis (Fig.16b and not shown). Since these organs have high proliferative activity, *hmg2* gene expression was also analysed in actively growing and growth-arrested NIH3T3 fibroblasts. Indeed *hmg2* is more expressed in proliferating than in growth-arrested NIH3T3 cells. This is in agreement with the evidence reported by Yoshida and co-workers that the level of hmg2 mRNA is enhanced in exponentially growing cells and cells transformed with various viral genes (Shirakawa and Yoshida, 1995; Yamazaki et al., 1995).

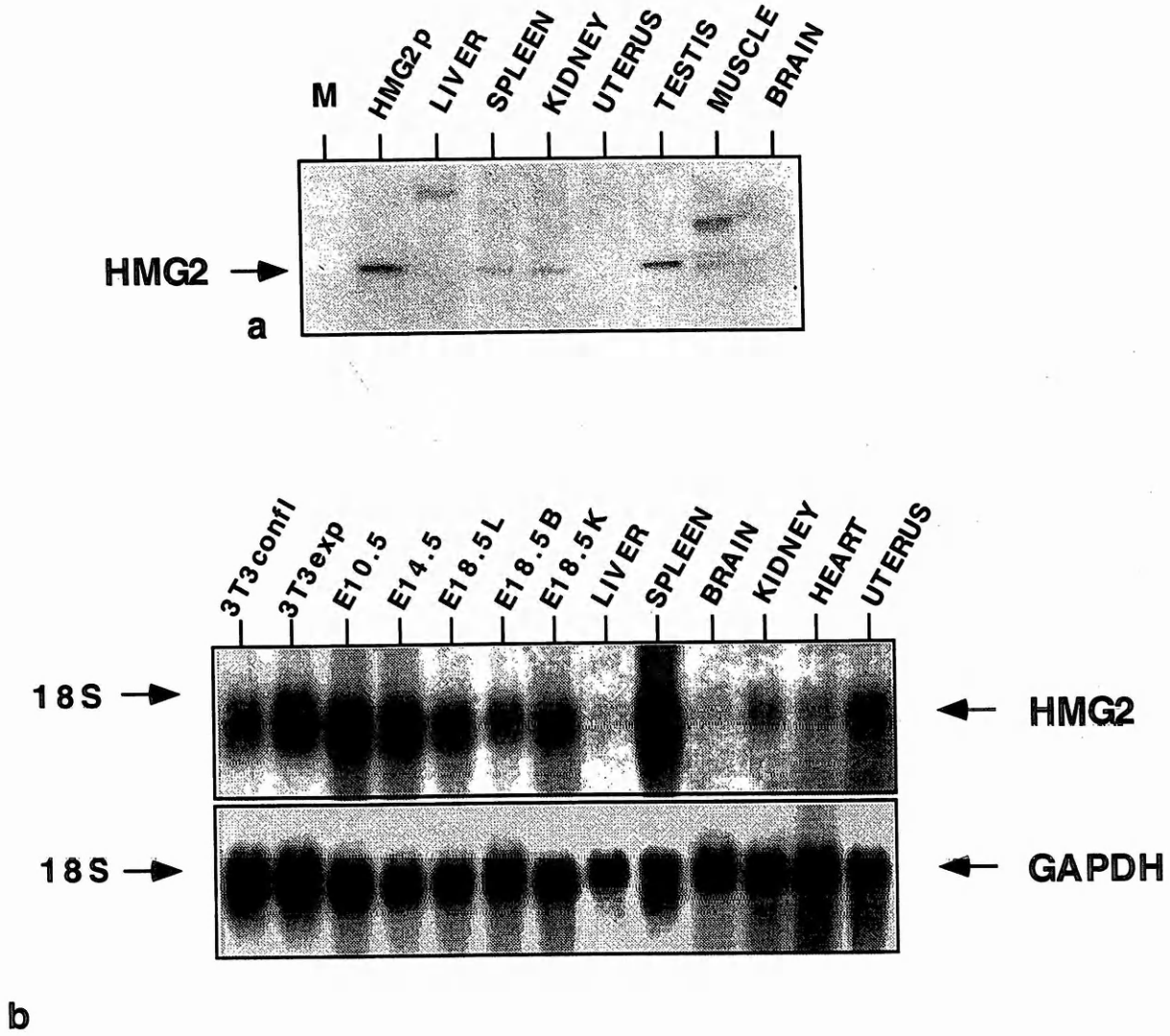


Fig.16 (a) HMG2 distribution analysis by Western blot. The same amount of tissue extracts were loaded. HMG2 is detectable in adult spleen, kidney, and it is very abundant in testis. Adult liver, brain, uterus, and muscle don't contain HMG2. M, marker; HMG2p, purified HMG2 protein used as positive control (100 ng).
(b) Northern blot analysis of *hmg2* gene expression: a single transcript of approximately 1 kb was detected using HMG2 cDNA as probe. The *hmg2* gene is highly expressed in exponentially growing NIH3T3 cells, while its expression is decreased in growth-arrested cells. The gene is also highly expressed in embryos at days 10.5, 14.5 and 18.5 liver (L), brain (B), and in adult spleen. Adult thymus and testis were also organs of high expression (not shown). Transcripts are low in adult uterus, kidney and heart and nearly undetectable in adult liver and brain.

In situ hybridisation was performed on development embryonic stages 10.5, 12.5 and 17.5. At E10.5 the gene is expressed widely, and especially in all sites of fast proliferation (Fig.17a). Fig.17b shows a frontal section at E12.5, where regional differentiation of the expression of *hmg2* is already evident. The strongest signal is present in the ventricular zone (VZ), which is a layer consisting exclusively of proliferating neuroepithelial cells. Expression is also high in the neural layer of the retina, while the expression of *hmg2* in the spinal cord forms a dorso-ventral gradient (arrow in Fig. 17b). At E17 (Fig.17c) *hmg2* is expressed in the ventricular zone and also in the differentiated cortex. Fig.18 shows in detail the anterior cerebral cortex shown in Fig.17c. *hmg2* signal is also present in the two parts that compose the transitional field (TF): the sub-ventricular zone, a zone whose nature remains poorly known, and the intermediate zone, to which differentiating cortical cells translocate before migrating to outer regions. The mesencephalon is also a site of strong signal (arrow 1 in Fig.17c). Thymus, lung, kidney cortex, the neural layer of the retina (Fig.17d), nasal dimps, olfactory bulb and the multilobular fat tissue are also regions of high expression, whereas lower but significant expression can be detected in the intestinal epithelium and in the hair follicles. At E17, *hmg2* is already expressed in the layer of cells that will form the External Granular Layer (EGL) of the cerebellum (arrow 2 in Fig.17c). Control hybridisation with the sense probe yielded a weak signal in liver; however, expression in liver at day 18,5 was demonstrated by Northern analysis (Fig.16b).

In the central nervous system (CNS), at postnatal day 4 (P4) *hmg2* is exclusively expressed in the external granular layer and in the hippocampus (Fig.19a). In a progressed stage (P17) a strong hybridisation signal is still present in the hippocampus and in the granular layer (GL) (arrows in Fig.19b); histological analysis of the slide revealed that also the Purkinje cells express the *hmg2* gene at high level.

Our analysis reveals that *hmg2* is not ubiquitously expressed, since it is absent in most of the adult tissues and since it is differentially expressed in the embryos, as shown

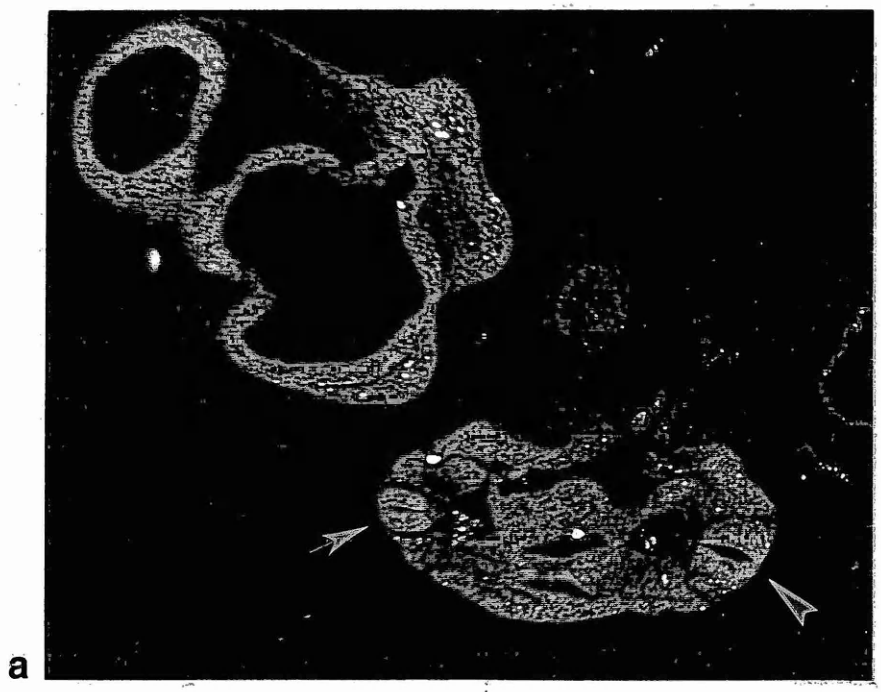


Fig.17a *In situ* hybridisation of a frontal section of E10.5. *hmg2* is widely expressed in all sites of fast proliferation. Number 1, 2, and 3 indicate respectively, the fourth ventricle, the third ventricle, and the telencephalic vesicle. Arrow head, indicates the neural tube in the caudal region of tail. Arrow indicates the neural tube in mid-tail region.

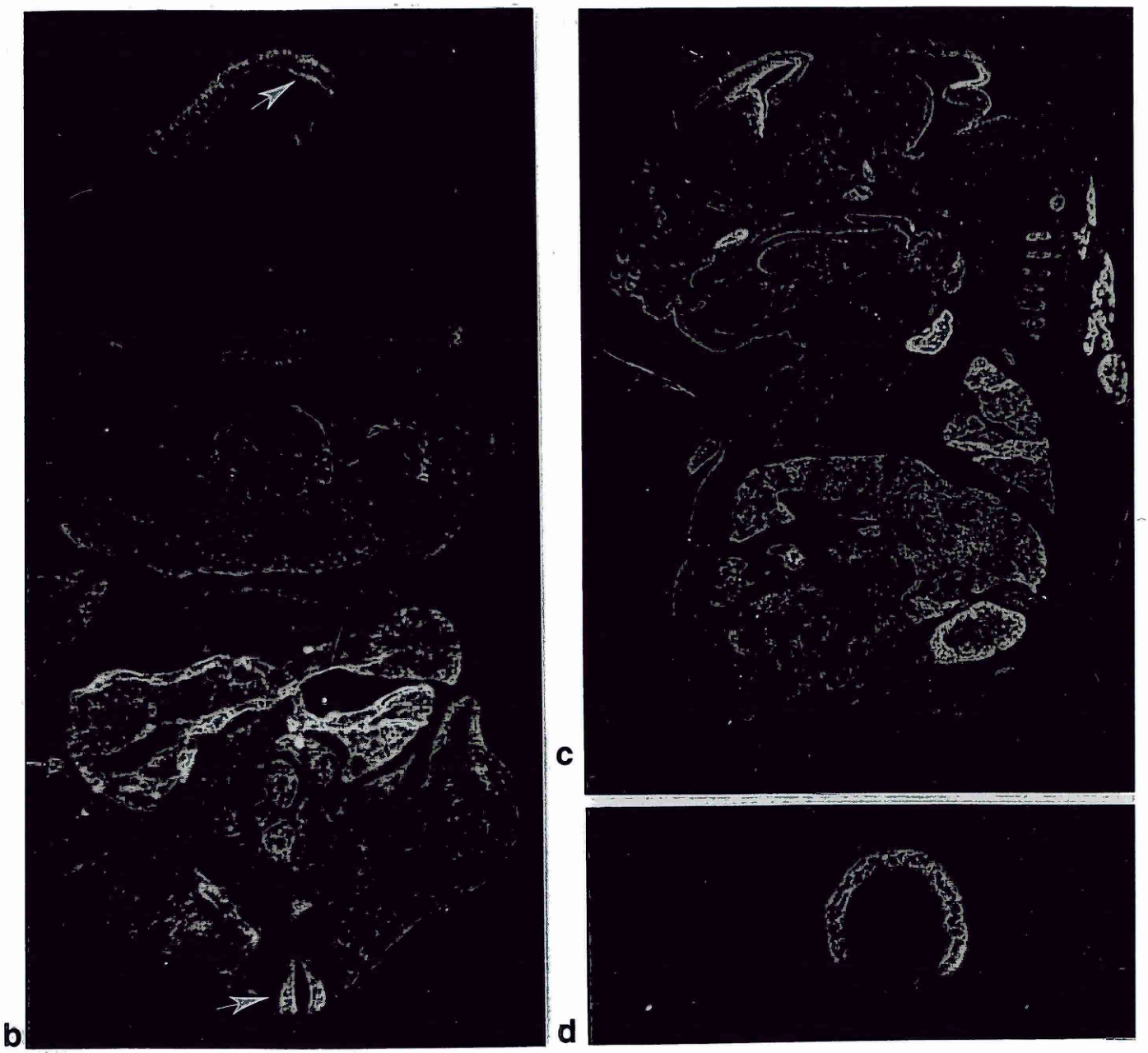


Fig.17 *In situ* hybridisation of different embryonic stages. b: Frontal section of E12.5. Strong signal is present in the ventricular zone (VZ) and in the spinal cord where it is possible to see a dorso-ventral gradient (arrows). c: Sagittal section of E17. *hmg2* is highly expressed in the ventricular zone and in the differentiated cortex (arrows). Strong signal is also present in the mesencephalon and in the layer of cells that will form the external granular layer of the cerebellum (arrows 1 and 2, respectively). Thymus, kidney cortex, and multilobular fat tissue are also sites of strong signal (arrow heads). d High resolution of the eye at E17: strong signal is present in the neural layer of the retina.

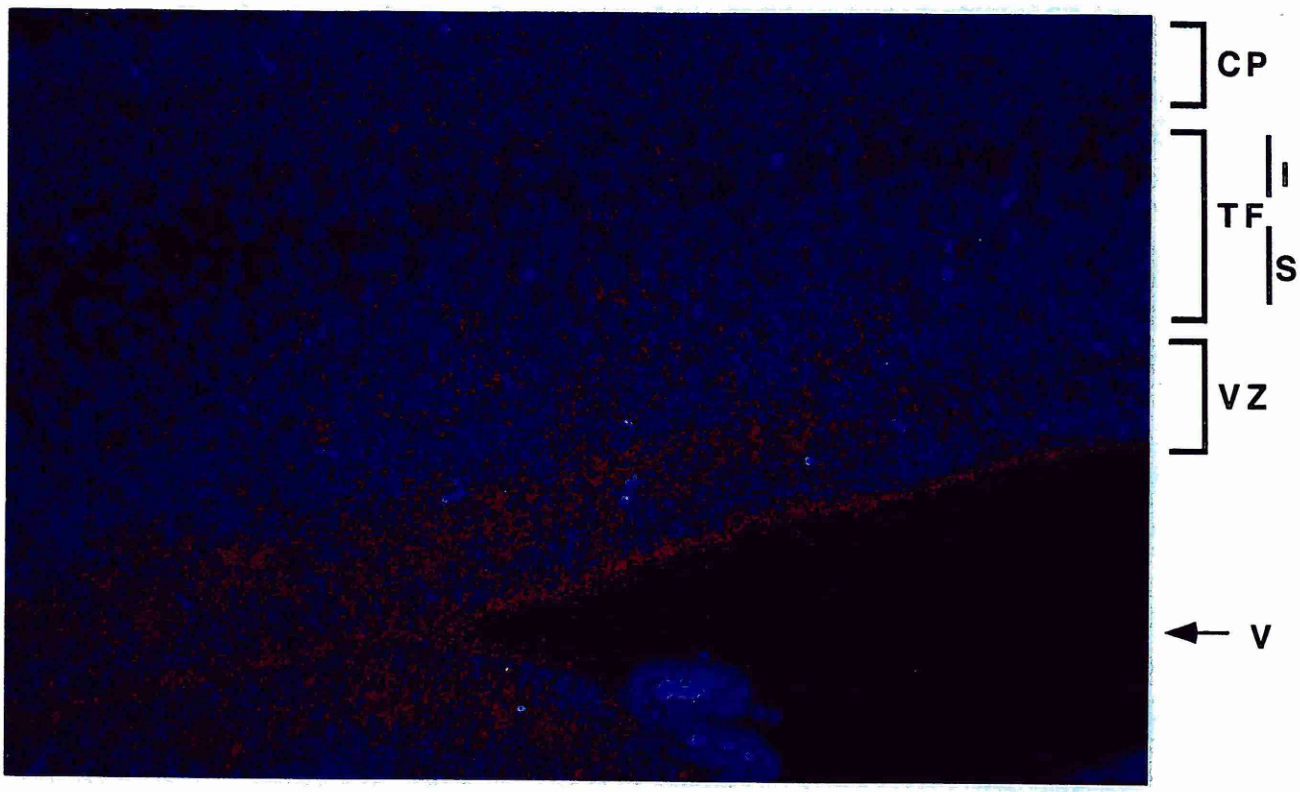


Fig.18 Higher magnification of the anterior cerebral cortex. Signal of hmg2 transcript is in red, while nuclei in the tissue were colored with Hoechst 33258, blue. V, ventricle; VZ, ventricular zone; TF, transitional field that is subdivided in two parts, the subventricular zone (SV) and the intermediate zone (IZ); CP, cortical plate. hmg2 signal is present in all these regions and in particular in the ventricular zone. Interestingly *hmg2* is also expressed in the intermediate zone, a region still poorly known.

Fig.19 a Expression of the *hmg2* gene in the cerebellum at Postnatal Day 4 (P4). Strong signal is present in the external granular layer (arrow). During maturation the cells that compose the external granular layer migrate toward the inside forming the internal granular layer. b At P17 strong signal is present in the Granular Layer (arrow). Arrow head shows that the transcript is also present in the dentate gyrus (arrow head).

by *in situ* hybridisation. Most of the sites that show strong signals are sites of active cell proliferation. Conversely, Fig.18 shows that *hmg2* is expressed in both proliferating and differentiated neurons that are present in the ventricular zone and in the cortical plate.

2.8 Production of recombinant clones of embryonic stem cells

After the characterisation of the gene the vector for gene targeting was constructed. Since *hmg2* is already expressed in embryonic stem cells (ES), I made a replacement construct that relies on the endogenous promoter of *hmg2* to express the positive selectable marker. With this construct the selectable gene should be expressed only in the case of homologous recombination. The vector, called pKO22, carries the selectable gene (the *sh-ble* gene that imparts zeocyn resistance) fused in frame to the first coding ATG of *hmg2,* and replacing the region that codes for HMG-boxA and part of the HMG-boxB (see Materials and Methods). In order to enrich for homologous recombinants, a negative selectable marker, the herpes simplex virus thymidine kinase gene (TK), was included outside the region of homology to the target gene. In the presence of the TK gene, the cells are sensitive to the gancyclovir. Therefore, homologous recombinants will be zeocyn resistant and gancyclovir resistant, whereas clones in which the construct integrated randomly will be zeocyn sensitive and gancyclovir sensitive.

I learnt to culture and to manipulate ES cells during a stage in the lab of my Second Supervisor, Dr. Hans Schöler, at the European Molecular Biological Laboratory (EMBL). I then set up the electroporation conditions of ES cells in our lab; at that time our Institute lacked competence in this field.

Two rounds of electroporations were done and the result is schematised in Fig.20. The first electroporation gave 68 clones resistant to zeocyn and sensitive to gancyclovir, the second gave 35 clones. Resistant clones were clonally isolated and screened by Southern blot analysis for the specific recombination allele (see Materials and Methods).

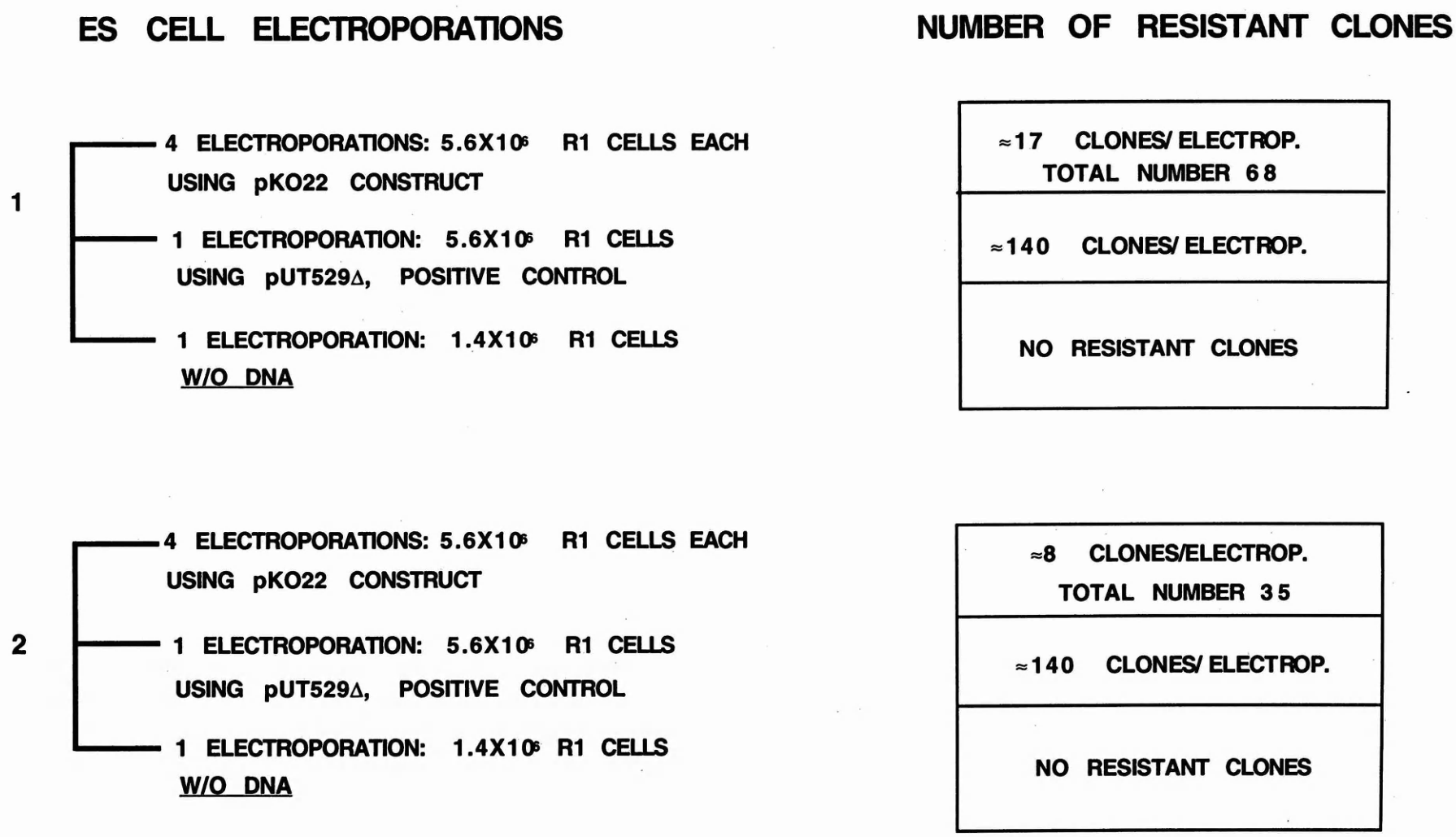


Fig.20 Two rounds of electroporation were done. For each electroporation, 2.2X10^7 R1 ES cells were transfected with pKO22 and gave 68 and 35 clones resistant to zeocyn and gancyclovir. 5.6X10^6 cells were also electroporated with pUT529Δ, that carries the *sh-ble* gene under the control of the TK promoter and Polyoma Enhancer, and gave 140 resistant clones. No clones were obtained in electroporations performed without DNA.

Total genomic DNA extracted from ES cells clones was digested with *Eco*RI restriction enzyme, that cuts outside the 5' homologous region comprised in pKO22 (Fig.21). The probe used to screen the clones spans a region of 700 bp near the *Eco*RI site and recognises a 4800 bp band in wild type DNA digested with *Eco*RI. After homologous recombination, two bands could be recognised by the probe: one corresponding to the wild type allele, and the other one corresponding to the recombinant allele. The latter is 7700 bp long because of presence of the selectable gene, that is 3600 bp long. As we can see in Fig.21, clones C5 and B3 are recombinant, and two bands with equal intensity are present. These clones were then amplified. Frozen samples were sent to the Biological Research Laboratory (BRL) for the blastocyst injection.

At BRL, the C5 and B3 clones were microinjected in 4 days old blastocysts from C57BL/6 donors, which were then transferred to a foster mother. The number of chimeric animals and the percentage of coat chimerism obtained is schematised in Fig. 22. The male animals with the highest percentage of chimerism were crossed with C57BL/6 females and tested for germ line transmission. All chimaeric males transmitted the *hmg2*$^-$ allele through the germ line.

2.9 Genotyping of the offspring

The littermates were first analysed by Southern blot. Total genomic DNA was extracted from tails and analysed using the same probe described in the previous chapter. In wild type mice a single band is present, in heterozygous mice two bands corresponding to the wild type and the knock-out alleles are detected, while in knock-out mice only the recombined allele is detected (Fig.23a).

The animals were routinely genotyped by PCR. Mice carrying the mutant allele were distinguished by the presence of the selectable gene, using a pair of oligonucleotides synthesised on the sequence of the *LacZ* gene. Knock-out mice were distinguished from

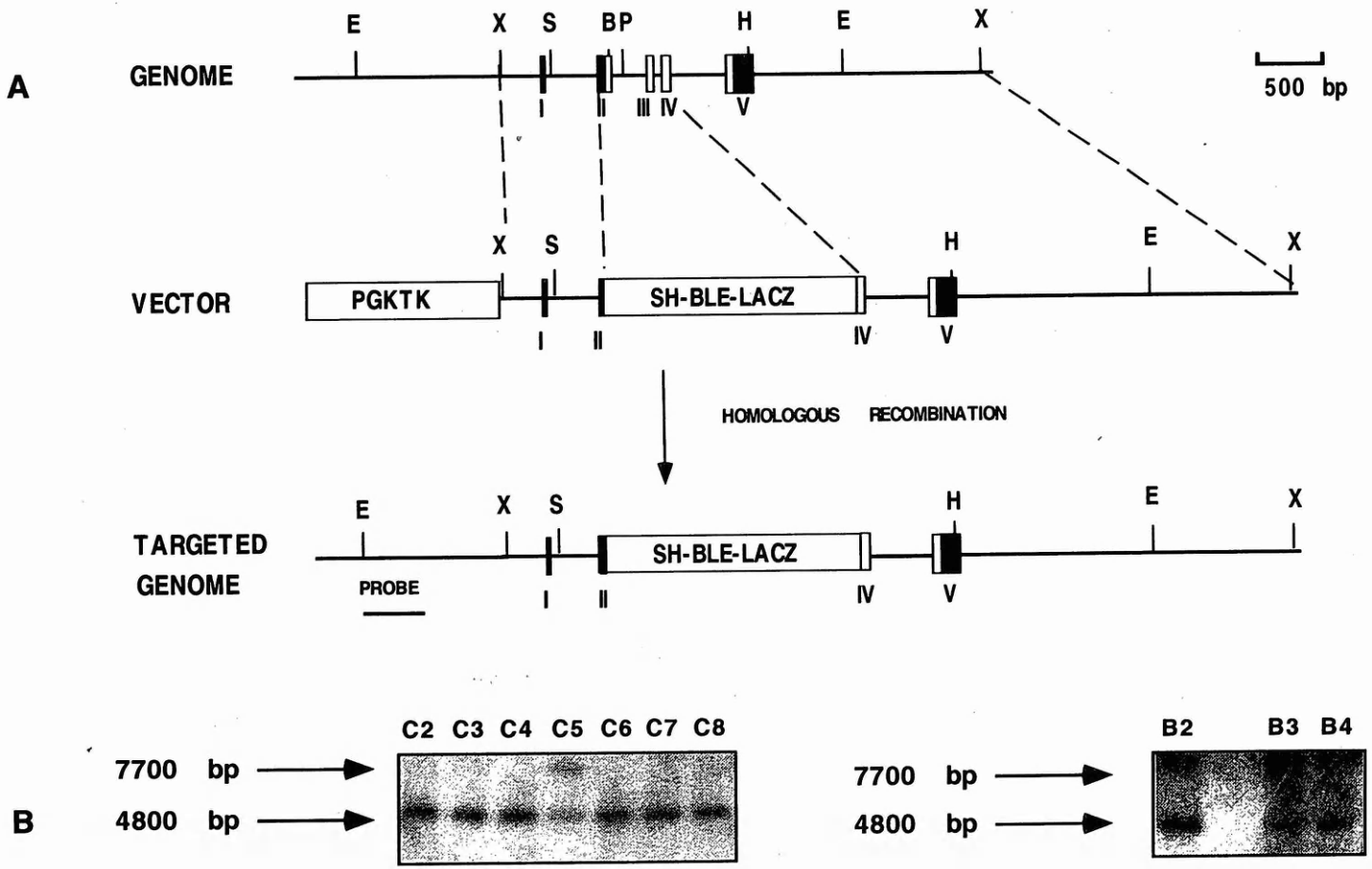


Fig.21 Replacement of *hmg2* by homologous recombination. (a) the genomic locus for *hmg2* contains five exons (vertical rectangles), the first of wich is untranslated. The targeting vector carries the *sh-ble* gene fused to *LacZ* and replacing exons 2, 3 and 4. The homologous region at 5' is 1 kb long, while the homologous region at 3' is 3.5 kb long. Outside the homologous region the targeting vector carries the thymidine kinase gene for the negative selection. (b) analysis of the ES resistant clones by Southern blot. Genomic DNA was cut with *Eco*RI restriction enzyme (E) and probed with a 700 bp fragment that maps to the right of EcoRI. This probe recognises the wild type allele (4800 bp) and the recombinant allele (7700 bp). C5 and B3 clones are recombinant.

RECIPIENT STRAIN	DONOR STRAIN	ES CELL LINE (CLONE)	TRANSFER DATE	NUMBER OF TRANSFERED BLASTOCYSTS	DATE OF BIRTH	NUMBER OF BORN ANIMALS	NUMBER OF CHIMERIC ANIMALS	SEX (% COAT CHIMERISM)
B6CBF1	B6	R1 C5	9-10-97	12	27-10-97	6	4	m (40, 3X100)
B6CBF1	B6	R1 B3	10-10-97	13	28-10-97	4	3	m (2X100) f (70)
B6CBF1	B6	R1 B3	10-10-97	13	28-10-97	5	4	m (2X100, 2X95)
B6CBF1	B6	R1 B3	10-10-97	14	28-10-97	7	2	m (10) f (30)
B6CBF1	B6	R1 B3	10-10-97	14	28-10-97	7	2	m (95) f (30)
B6CBF1	B6	R1 B3	10-10-97	14	28-10-97	9	3	m (2X90) f(10)
B6CBF1	B6	R1 B3	10-10-97	16	28-10-97	9	4	m (3X90) f (80)
			TOTAL	96		47	22	17m/5f

Fig.22 Blastocyst injection of the two positive clones I obtained after electroporation of R1 ES clones B3 and C5. The blastocyst injections were done in the Biological Research Laboratories (BRL). 4 days old blastocysts from C57BL/6 were microinjected with B3 and C5 clones, respectively. The male mice with higher percentage of chimerism were mated with C57BL/6 females and analysed for germline transmission.

the heterozygous by the absence of the intron 3, using PCR with oligonucleotides flanking intron 3 (Fig.23b).

Fig.23c1 shows the Western blot analysis of protein extracts from adult testes. In lane 1, an extract of NIH3T3 which were transiently transfected with pHMG2 was run as a positive control. An antibody against HMG2 recognised a clear band in the positive control and in testes extracts from wild type and heterozygous mice. In +/- mice the quantity of HMG2 decreases while the slower migrating band indicated by two asterisks in Fig.23c1, appears more intense. In -/- extracts the HMG2 band disappears and the upper band is augmented. At the moment we ignore the nature of the slower migrating band, which is present in all the extracts, wild type or otherwise, that we analysed. Since we have only this antibody against HMG2, it is possible that the slower migrating band represents a protein that fortuitously cross-reacts with our antibody. Alternatively, it could be another HMG protein produced to compensate the lack of HMG2. So far have not found proteins of the HMG1/2 family of the size corresponding to the slower migrating band. The weak band present in the lane of -/- extract is probably HMG1 (Fig.23c1, one asterisk): in fact, the same filter incubated with an antibody against HMG1 (Fig.23c2) shows that HMG1 migrates at the same level.

2.10 Phenotypic analysis of HMG2-deficient mice

Heterozygous mice were normal and produced offspring. Knock-out mice as well had no obvious phenotypic differences from wild type. The external appearance, gait behaviour, weight, fertility, and life span of the homozygotes appeared normal. Organs from wild type and knock-out mice were analysed: we looked in particular at the size and morphology of the organs where HMG2 is highly expressed during embryogenesis. However, brain, thymus and testes of several knock-out mice showed no macroscopic defects. In spite of the interaction of HMG2 with HOX proteins, homozygote mice showed no homeotic defects; in particular, vertebrae and ribs were normal.

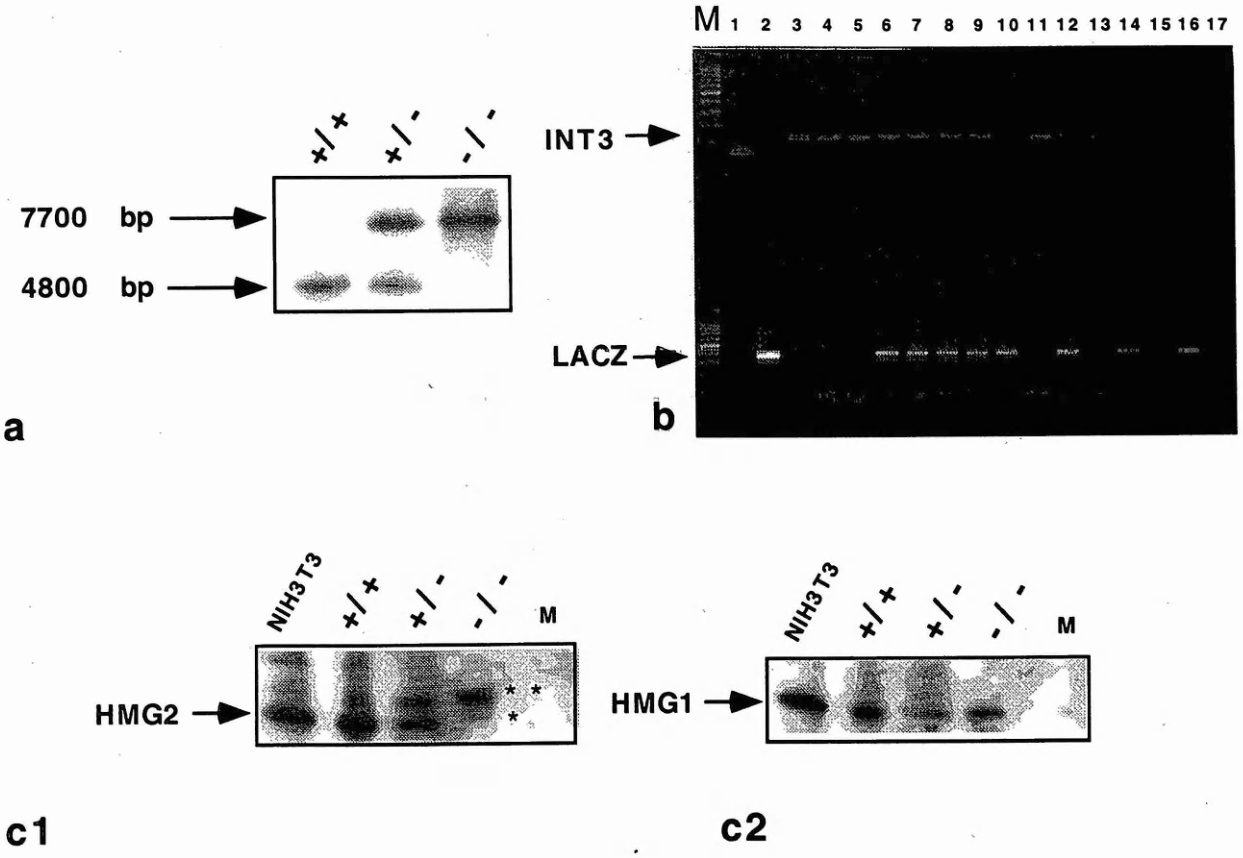


Fig.23 a Southern blot analysis of genomic DNA extracted from KO and wild type mice a single band of 4800 bp is detectable. In heterozygous mice two bands are detectable. In homozygous mice only the recombined allele is present. b Mice were routinely genotyped by PCR. Heterozygous and knock-out mice were distinguished by the presence of the selectable gene (*sh-ble* fused to *lacZ*), using a couple of primers synthetised on the sequence of *lacZ* gene (See Materials and Methods), as shown at bottom of the gel. Knock-out mice were distinguished from heterozygous by the absence of intron 3, as shown in the upper part of the gel. Mice 10 and 16 are knock-out. c1 Western blot analysis using an antibody against HMG2. In extracts from NIH3T3 cell and wild type testis the HMG2 band is present, in +/- HMG2 is also present and the slower migrating band present already in wild type increase the intensity (two asterisks). In -/-, HMG2 disappears but the upper band is augmented. The slower migrating band could be due to cross-reaction with HMG2 antibody or could be an unknown HMG protein. c2 The same filter was incubated with an antibody against HMG1. HMG1 migrates at the level of the weak band, indicated by an single asterisk in c1. HMG1 protein was weakly recognised by the antiHMG2 antibody. This result excludes that the slower migrating band could be HMG1.

Since HMG2 is expressed at high level in the hippocampus and in the external granular layer of the cerebellum during their development, *hmg2*$^{-/-}$ homozygotes were subjected to behavioural tests in the lab of Dr. D. Wolfer in the University of Zurich. No difference from wild type was found in the Morris water maze test, that is sensitive to hippocampal lesions (Stewart, 1993; Stewart and Wong, 1993), and in the two-way active avoidance test, a measure of emotional responses (Clincke and Wauquier, 1979). These results suggest that *hmg2*$^{-/-}$ homozygotes have no gross alteration of cognitive functions or coordination.

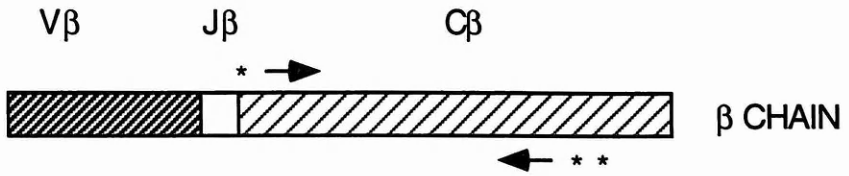
2.11 The TCR repertoire is not affected in HMG2-deficient mice

HMG1 and HMG2 are involved in V(D)J recombination, which generates both the diversity of antibodies produced by B cells and TCRs expressed on T cells. RAG1 and RAG2, which were recently demonstrated to interact with HMG1 and 2, play a crucial role in this process (Aidinis et al., 1999). We therefore investigated whether V(D)J recombination was altered in -/- mice.

When V(D)J recombination occurs normally the TCR repertoire is polyclonal; the presence of oligoclonal T cells is consequently a symptom of an inefficient V(D)J recombination. We analysed the variable ß-chain (TCRBV) usage and we characterised the TCRBV clonality of individual TCRBV segments in knock-out mice with the heteroduplex technique (Wack et al., 1996). This technique is based on the amplification by PCR of different variable families of TCRBV: the denaturation of the PCR products obtained, followed by a renaturation at a permissive temperature, allows the formation of heteroduplexes and homoduplexes. Electrophoretic migration in native polyacrylamide gels allows the discrimination of bands corresponding to homoduplexes and heteroduplexes.

The TCR repertoire of three knock-out and one wild type mice was analysed. Total RNA extracted from peripheral blood T lymphocytes was retrotranscribed and the

quality of the cDNA was analysed by PCR: a 450 bp long fragment of the TCR ß chain was amplified using a couple of primers synthesised on exon 1 and 3, respectively, of the constant ß chain (Cß) (upCßex1 and dwCßex3) (schematic representation of primers and PCR results are shown in Fig.24). The cDNA of T lymphocytes was then analysed for seven different variable families (see Materials and Methods): Vß 2, Vß5.1, Vß5.2.3, Vß6, Vß8.1, Vß8.2 and Vß8.3. Seven different PCRs were done for each mouse to be analysed. In each PCR reaction we used the same downstream primer (dwCßex1) annealing to the constant ß fragment and a specific primer annealing to the variable region of the ß chain. PCR products were denatured at 94°C for 5 minutes, and renatured at 50°C for 60 minutes. The fast passage from 94°C to 50°C allows the random matching of single strand DNA molecules. The annealing of non-complementary molecules deriving from different clones of the same Vß family produces heteroduplexes. PCR products were run on PAGE and then transferred onto a nylon filter and hybridised with an oligonucleotide that maps on exon 1 of the ß-chain, dwex1Cß (Fig.25a). In both wild type and knock-out mice the TCR repertoire of the tested Vß families was found to be polyclonal. The polyclonality of T cells is visible as a smear (Fig. 25b). In the case of oligoclonality the amplification of one or few specific bands is expected (Wack et al., 1996). Therefore, V(D)J recombination occurs normally in -/- mice, in spite of the demonstration of a crucial role of HMG2 in the first step of the V(D)J recombination (Agrawal et al., 1998; Agrawal and Schatz, 1997; Aidinis et al., 1999). In any case, we are not surprised by the absence of effects on the TCR repertoire in *hmg2* $^{-/-}$ mice since it was demonstrated that *in vitro* HMG1 and HMG2 play the same role during V(D)J recombination, and it is highly possible that HMG1 can vicariate HMG2.



* upCβex1 5'-GAAATGTGACTCCACCCAAG-3'

** dwCβex3 5'-AGGATCTCATAGAGGATGGT-3'

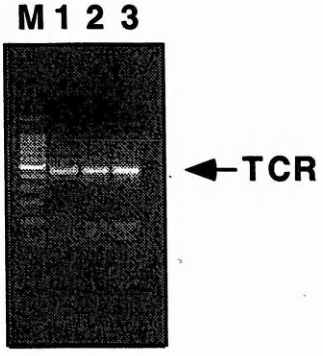


Fig.24. Schematic representation of oligonucleotides used to amplify a 450 bp long fragment on the costant β chain of the TCR. T lymphocytes were isolated from peripheral blood, RNA was extracted and retrotranscribed. Quality of cDNA was tested using a pair of oligonucleotides synthesised on exon 1 (mCβex1up) and on exon 3 (mCβex3dw) of the constant segment of the β chain of TCR. The arrow indicates the amplified fragment in three of four mice analysed. M, marker 100 bp ladder from Promega. PCR was done on -/- mice number 1, 2, and 3. The PCR result on the +/+ mouse is not shown in this figure.

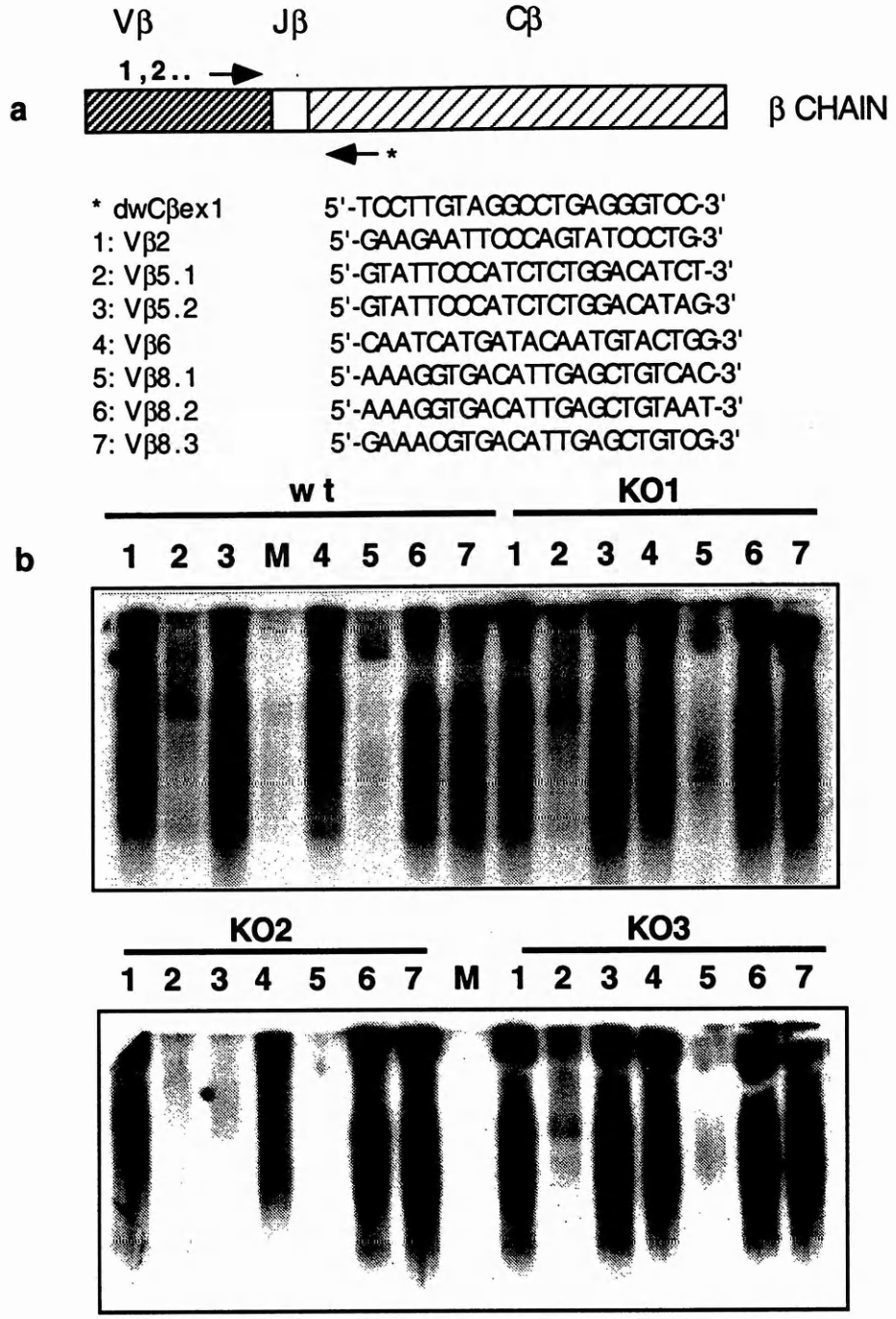


Fig.25a Strategy used to amplify seven Vβ families of TCR b chain and sequences of the oligonucleotides. Fig.25b Seven Vβ families were analysed in three knock-out mice and a wild type mouse. Each number corresponds to the family analysed, for example 1 is for the Vβ2 family. The lane of the marker is indicated by M. The presence of smear indicates the formation of heteroduplexes during the renaturation of PCR products. No major differences are visible between wild type and knock-out mice .

2.12 *hmg2* knock-out male mice have reduced fertility

The major defect we found in -/- mice is a partial male infertility. Two continuous breeding studies were done for 3 and 2 months, respectively. On the first study three knock-out mice were mated with wild type females. As schematised in Fig.26, the A, B and C -/- males gave no littermates. A and B derive from chimeras obtained by the ES cells clone B3, while C derives from the C5 clone. Therefore, the same phenotype was obtained by the two recombinant clones. During the same time wild type D, E and F mice gave at least two littermates. It is interesting to note that knock-out females are apparently normal. During a second continuous breeding study A, B and C gave no littermates again, while other four animals analysed gave littermates of variable size. Two gave normal littermates while the other two gave litters drastically reduced in size. Then two brothers (number 0 and number 3) were mated to different females and vaginal plugs were detected. Of the two females mated with male number 0, one gave no littermate while the other one gave only one pup. Of the three females mated with male number 3, two generated littermates of three pups each and the other one a littermate with 8 pups.

2.13 Histological defects are present in seminiferous tubules of -/- mice

To investigate the partial male infertility, we looked for the size and the histology of the testis in different knock-out mice.

No differences were found in the weight and size of the testis between wild type and knock-out. Therefore they were paraffin-embedded and cut in thin sections for histological analysis. Fig.27a shows a section of a 163 days old -/- mouse testis. Some seminiferous tubules are completely degenerated, with few or no germ cells remaining, and some tubules are apparently normal and able to produce sperms. Around 15% of the tubules are degenerated and the degree of degeneration is variable. Fig.27b shows a higher magnification of an apparently normal seminiferous tubule. Nevertheless, it

FIRST CONTINUOUS BREEDING STUDY (THREE MONTHS)

males			females	
A	- / -	X	+/+	NO LITTERS
B	- / -	X	+/+	NO LITTERS
C	- / -	X	+/+	NO LITTERS
D	+/+	X	- / -	AT LEAST TWO LITTERS
E	+/+	X	+/+	AT LEAST TWO LITTERS
F	+/+	X	+/+	AT LEAST TWO LITTERS

SECOND CONTINUOUS BREEDING STUDY (TWO MONTHS)

A	- / -	X	+/+	NO LITTERS
B	- / -	X	+/+	NO LITTERS
C	- / -	X	+/+	NO LITTERS
D	+/+	X	- / -	ONE LITTERS
E	+/+	X	+/+	ONE LITTERS
F	+/+	X	+/+	ONE LITTERS
G	- / -	X	+/+	ONE LITTER (2 PUPS)
H	- / -	X	+/+	ONE LITTER(12 PUPS)
I	- / -	X	+/+	ONE LITTER (10 PUPS)
L	- / -	X	+/+	ONE LITTER (2 PUPS)

Fig.26 Continuous breeding studies. KO males A, B and C gave no littermates in both runs. KO males G and L are partially sterile, while KO males H and I are fertile.

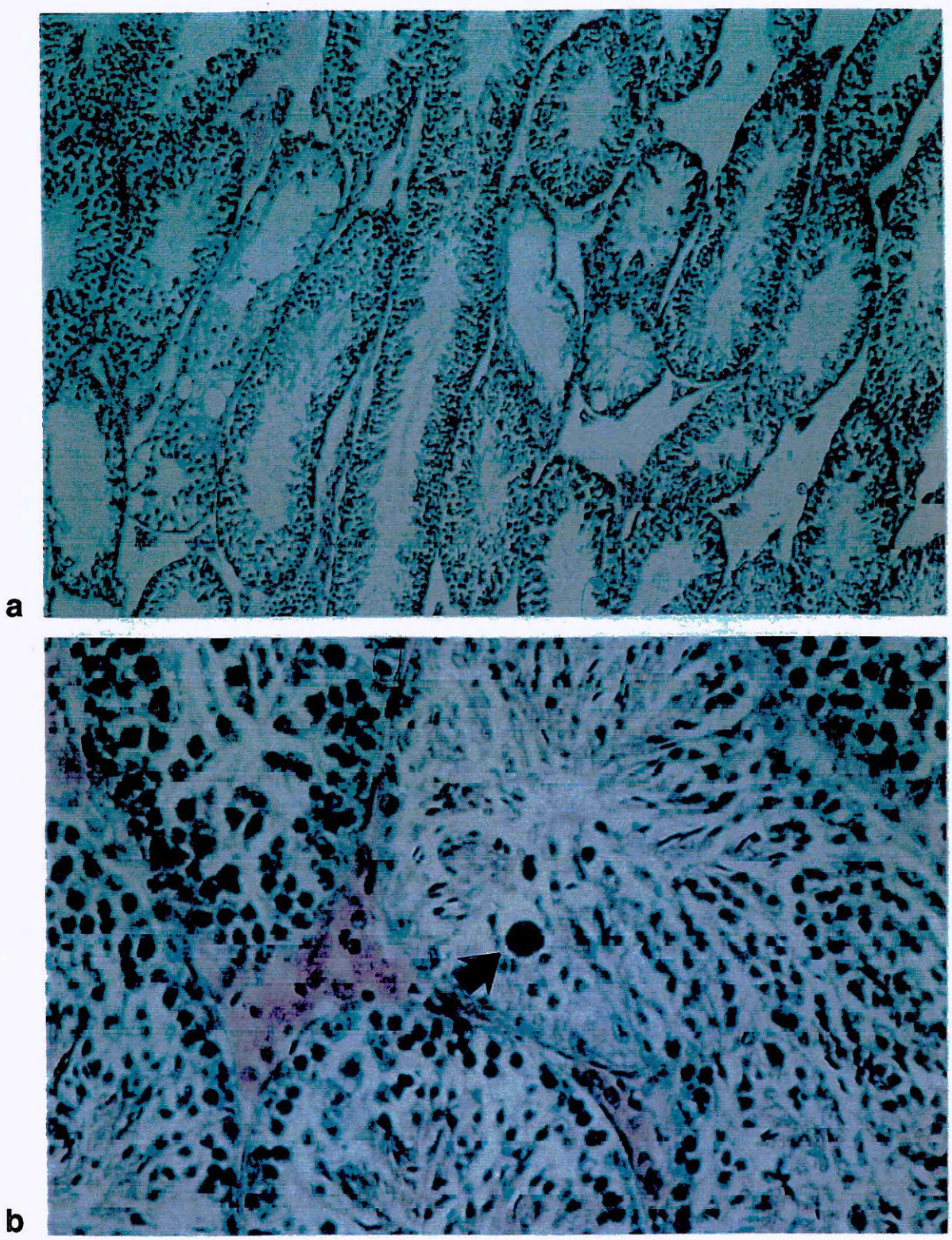


Fig.27 (a) Ematoxylin-eosin staining of a testis section from a 163 days old mouse. Arrows indicate seminiferous tubules with different degrees of degeneration. Most of the tubules are normal. (b) higher magnification of an apparently normal seminiferous tubule, nevertheless it contains a multinucleated cell with pyknotic nucleus. We found this kind of cell only in knock-out sections and associated with the meiotic wall of the tubule.

contains multinucleated cells with pyknotic nuclei. Such cells were found only in knock-out mice. Fig.28a shows an example of a seminiferous tubule where the degeneration is not complete. Germ cells are still present, but the cytoplasms of Sertoli cells contain large vacuoles. Fig.28b shows a higher magnification of a severely damaged tubule that contains vacuolated Sertoli cells. The organisation of germ cells is completely lost and sperms are not produced. On the contrary, spermatozoa are produced in the tubule in shown in Fig.28a. The defects described were found both in completely sterile and fertile -/- mice. The latter have a less severe defect, suggesting that the phenotype has a variable penetrance.

Fig.29 shows a section of the testis of a young -/- mouse where degenerated seminiferous tubules are not present. The only detectable defect is the presence of the multinucleated cells with pyknotic nuclei. Testes from animals younger than one month show no apparent defects. Thereafter, tubules undergo distension with a progressive increase in dysmorphogenesis resulting from atrophy. The severity of the defect is therefore dependent on the age of the mouse.

2.14 In apparently normal seminiferous tubules of HMG2-deficient mice the number of apoptotic cells is increased

The histology of -/- tubules revealed the presence of severe defects: some tubules were completely degenerated and others lost the classical cellular organisation. Moreover, the presence of multinucleated cells with condensed nuclei that appeared fragmented and apoptotic moved us to verify whether in -/- tubules programmed cell death was triggered. Apoptosis was detected by *in situ* terminal deoxytransferase-mediated dUTP nick end labelling (TUNEL) in testis sections from 160 days old wild type and knock-out mice. In wild type tubules a very low number of apoptotic cells was found; most tubules contain no apoptotic cell (Fig.30a). On the contrary, in knock-out sections we found from 1 to 16 apoptotic cells per tubule: all tubules contain at least one apoptotic cell (Fig.30b). The

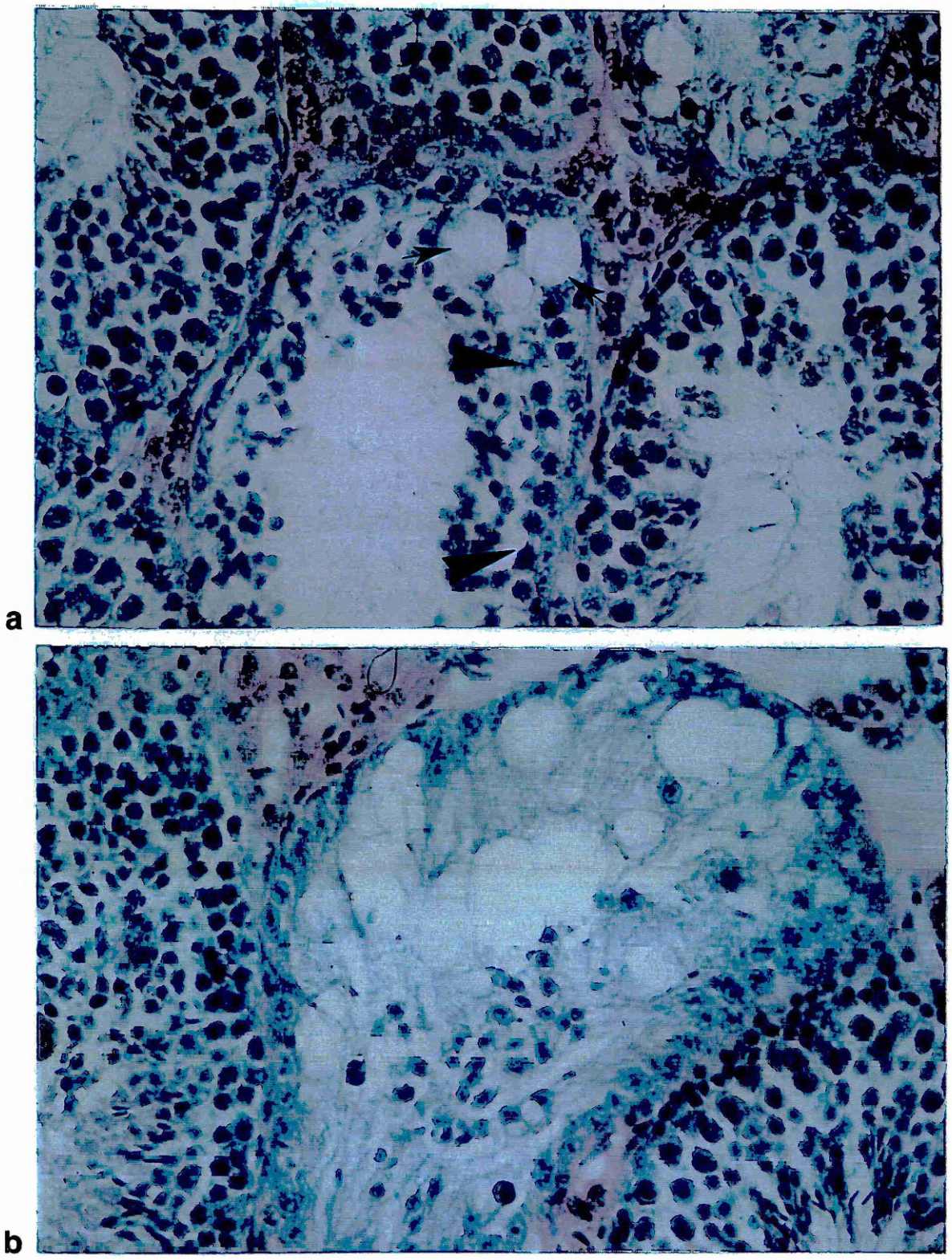


Fig.28 (a) Higher magnification of a seminiferous tubule were degeneration is already evident. Germinal cells are still present (arrow head). Vacuoles in the cytoplasm are present (arrows) and sperms are produced. (b) Example of a tubule severly damaged. Nuclei of Sertoli cells are indicated by arrows.

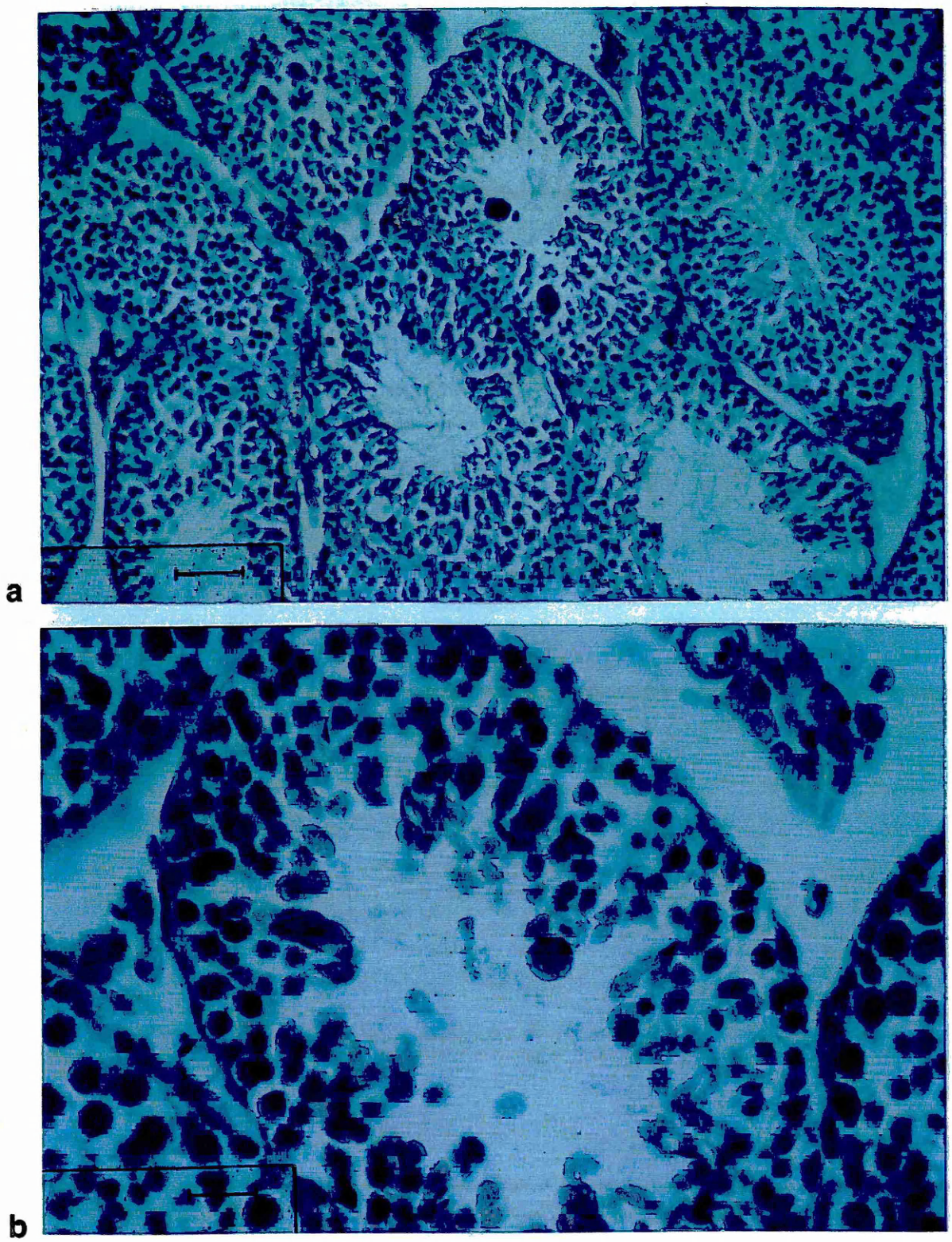


Fig29 (a) Ematoxylin-eosin staining of a testis section from a 60 days old mouse. The seminiferous tubules are apparently normal. (b) Higher magnification of a seminiferous tubule containing the multinucleated cell we described to be frequently present in tubules of older mice. In young mice the frequence of these cells is very low. No vacuoles and cellular disorganisation were found.

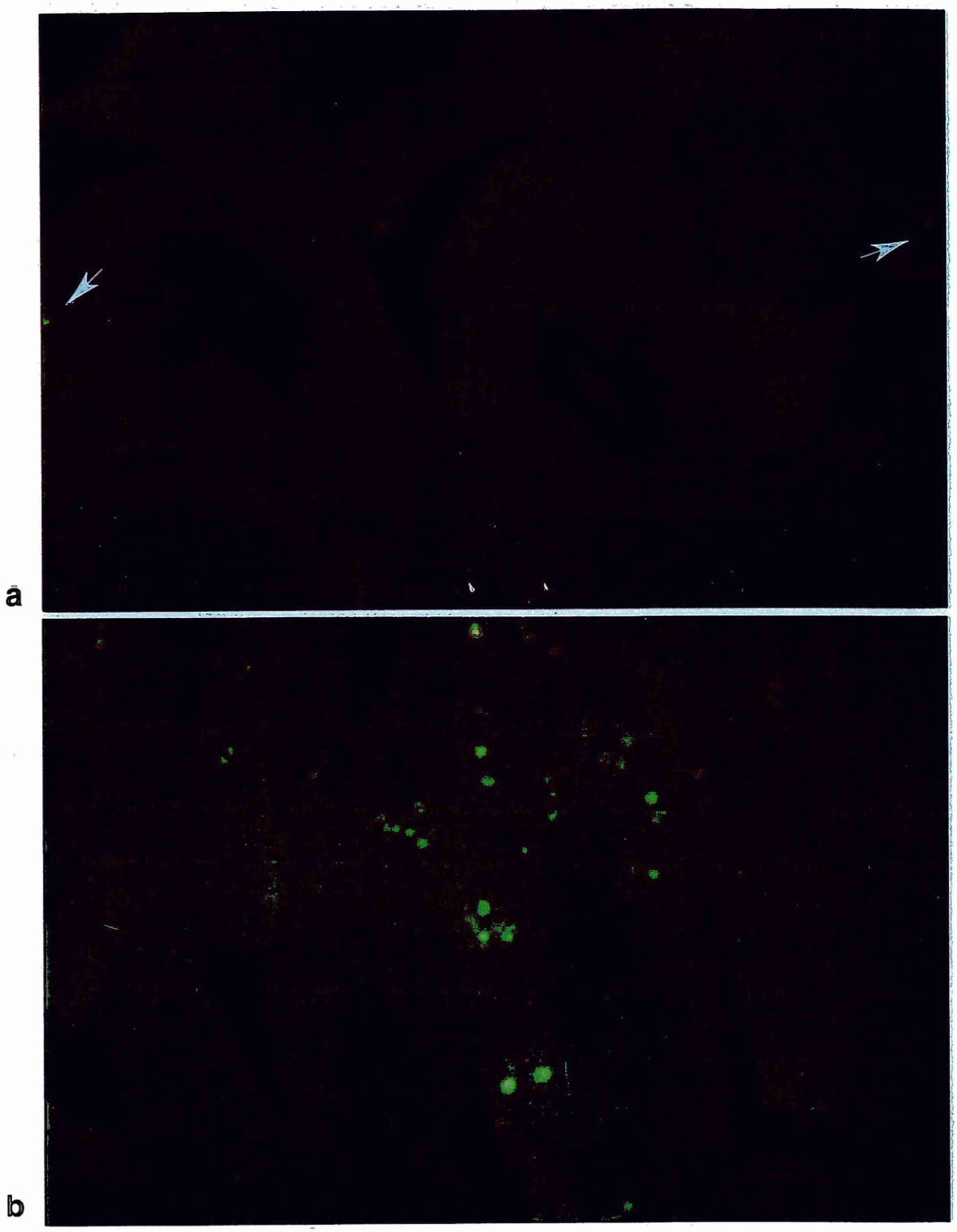


Fig.30 (a) TUNEL-staining on wild type section. Two apoptotic cells are present (arrows). (b) Several apoptotic cells were found in knock-out sections. All tubules contain apoptotic cells.

number of apoptotic cells was counted in randomly chosen seminiferous tubules, 30 from wild type and 30 from knock-out mice. In -/- sections we counted a mean of 4 apoptotic cells per tubule, against less than 1 per wild type tubule. In general, apoptotic cells were found both in apparently normal and in degenerated tubules. Therefore, also tubules without any histological defects show subtle abnormalities.

2.15 Distribution of HMG2 during spermatogenesis

Seminiferous tubules are composed of a stratified epithelium of germ cells surrounding a lumen. As in other stratified epithelia, cells become increasingly more differentiated as they move away from the basement membrane. Spermatogonia, the proliferating cells, are located near the basement, spermatocytes are more medially located, and spermatids are found nearest to the lumen. As they mature, spermatozoa are released into the lumen. The different cell types of the seminiferous tubule are easily distinguished by their position in the tubule, their size, the shape of their nucleus, the chromatin density and the ratio between cytoplasm and nucleus. An example of different cell types in the tubule is schematised in Fig.31.

Contradictory data were published on the distribution and quantitation of HMG2 protein in the different cell types of seminiferous tubule. Seyedin et al. (1979) demonstrated that HMG2 is present at high level in all tissues with high proliferative activity as adult thymus, spleen and testis. Consequently, they hypothesised that very high levels of HMG2 is associated with proliferative activity. On the contrary, Bucci et al. (1984, 1985) demonstrated that HMG2 levels are low in spermatogonia, the only proliferating cells of the testis, and in preleptotene spermatocytes. In pachytene spermatocytes and early spermatids the levels HMG2 are very high, while the protein is not detectable in late spermatids and in spermatozoa.

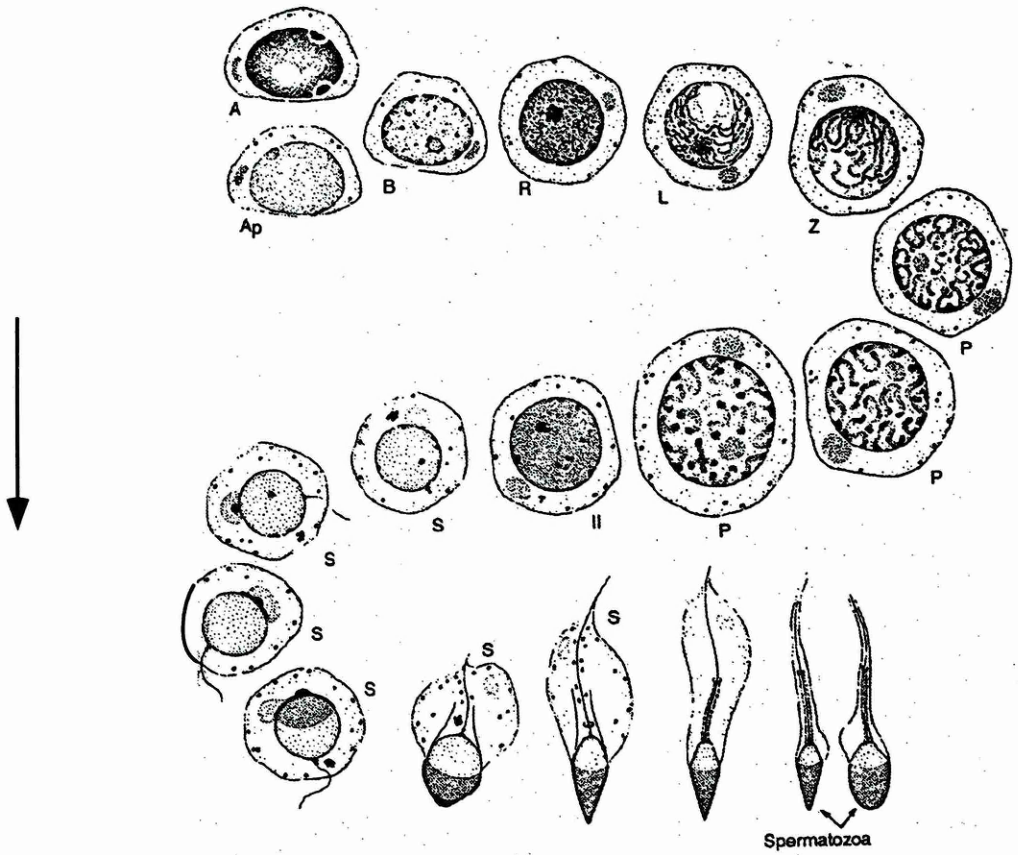


Fig.31 Shape of male germ cells in the mouse seminiferous tubule. A, represents Type A spermatogonia; AP, Type A spermatogonia in prophase; B, Type B spermatogonia; R, resting primary spermatocytes; L, leptotene primary spermatocytes; Z, zygotene primary spermatocytes; P, pachytene primary spermatocytes; II, secondary spermatocytes; S, spermatids in different stages of spermatogenesis (from Clermont 1960, *Fert. Steril.* 11, 563)

We analysed which cell types contain abundant HMG2 protein using *in situ* techniques: immunohistochemistry to analyse the distribution of the protein and *in situ* hybridisation to analyse the expression of the *hmg2* gene.

For immunohistochemistry, sections from wild type testis were incubated with HMG2 antibody from PharMingen (see Materials and Methods). Sections were developed with 3-amino-9-ethyl carbazole (AEC) chromogen, that produces a red precipitate. To distinguish the different cell types, the sections were mounted with an aqueous mounting medium containing 4',6-diamidino-2-phenylindole (DAPI). DAPI binds DNA and the cells of the seminiferous tubule can be distinguished by their characteristic chromatin organisation, as shown in the Fig.31. Intense red colour was found in pachytene spermatocytes and in round spermatids (open arrows and closed arrows in Fig. 32a). Little signal was detected in spermatogonia and in preleptotene spermatocytes. No signal was found in late spermatids. The counterstain with DAPI shows clearly cells that are tightly associated to the basement membrane of the tubule and form a circle round pachytene spermatocytes. Their chromatin organisation indicates that they are spermatogonia and preleptotene spermatocytes, as shown by arrow heads in Fig.32b.

The data obtained by immunohistochemistry were confirmed by the analysis of ß-galactosidase activity (ß-gal). In heterozygous and in knock-out mice *Lac-Z* gene expression reflects the expression pattern of the *hmg2* gene, since the *sh-ble:lac-Z* coding sequence is fused in frame to the ATG of *hmg2*. Testes from heterozygous and wild type mice were dissected, fixed in fresh 4% paraformaldehyde and incubated overnight at room temperature with X-GAL substrate. Testes were then paraffin embedded, cut in 7 µm slices and analysed by light microscopy. In +/- testes a pale blue colour was found in the cytoplasm of spermatogonia and preleptotene spermatocytes. On the contrary, intense blue colour was found in the cytoplasm of spermatocytes and in the cytoplasm of early

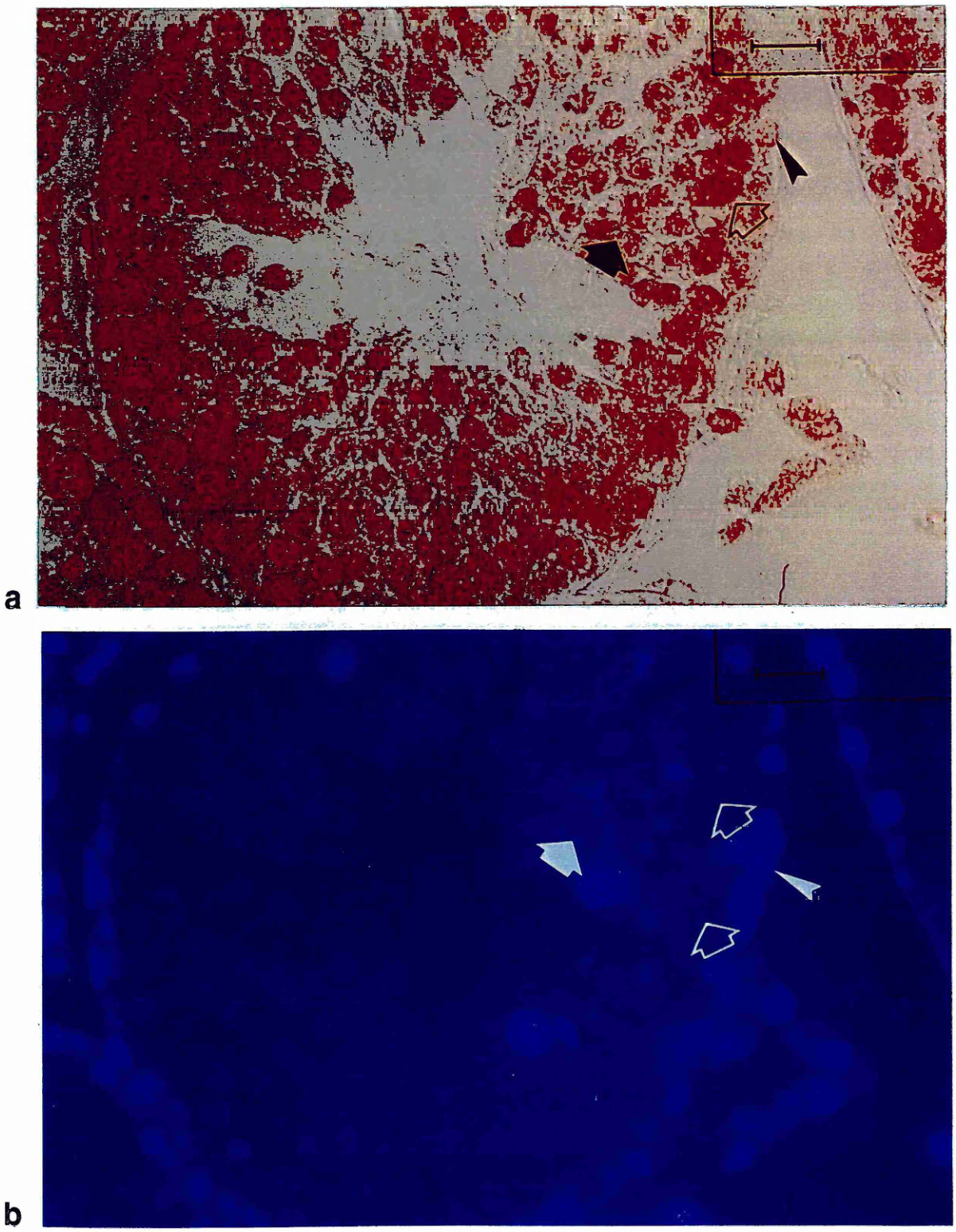


Fig.32 (a) Immunohistochemistry of a testis section from 60 days old wild type mouse. HMG2 protein is present in pachytene spermatocytes and in round spermatids (open and filled arrows, respectively). (b) The same tubule in 32a was counter stained with DAPI and photographed at fluorescence microscope. Arrow head shows spermatogonia, open arrow indicates the position of spermatocytes while filled arrow the position of round spermatids.

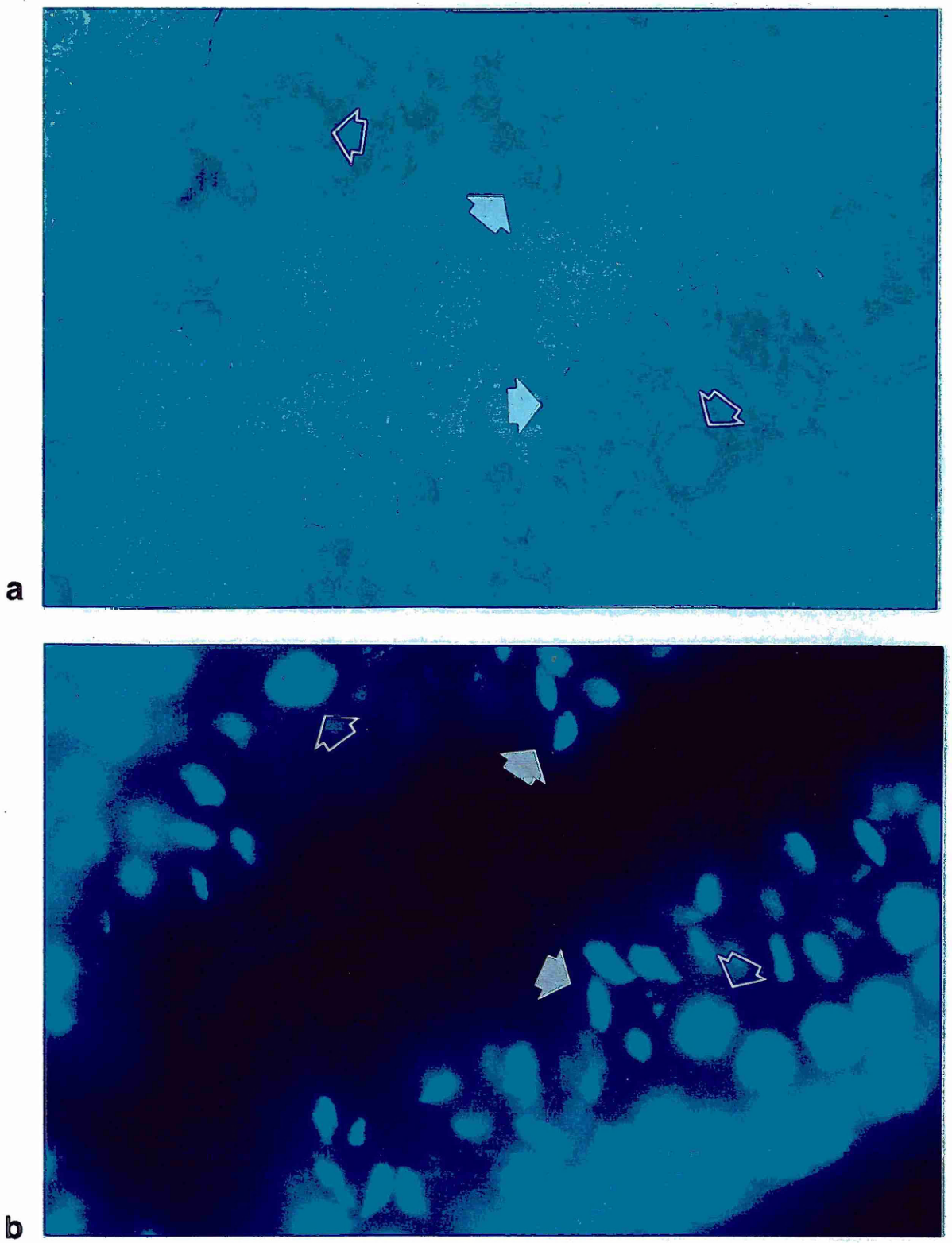


Fig.33 (a) β-gal staining of testis section of a 60 days old heterozygous mouse. Intense blue staining is on the cytoplasm of spermatocytes (open arrow) and on the cytoplasm of round spermatids (not shown). Blue color is also present in elongating spermatids (arrow). (b) counter stained section: the same kind of arrows used in 33a were used to indicate the different cell types.

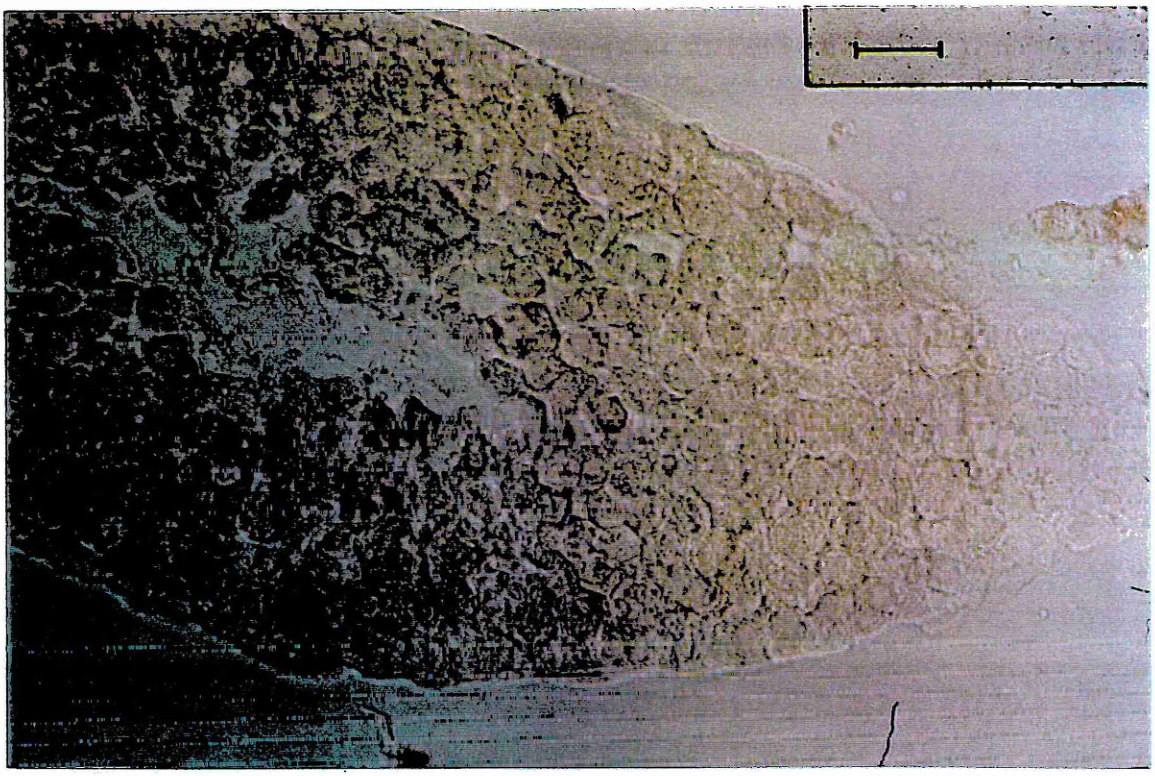


Fig.33 (c) β-gal staining of +/+ testis sections of a 60 days old wild type mouse. As expected no color was found.

and late spermatids (arrows in Fig.33a), as confirmed by DAPI staining in Fig.33b. In +/+ sections ß-gal activity was not detected (Fig.33c).

Levels of HMG2 protein and ß-gal activity found in the nucleus of pachytene spermatocytes and round spermatids are in agreement. However, ß-gal activity was detected at high level in late spermatids, where no HMG2 appears to be present. The most plausible explanation is specific HMG2 degradation, whereas ß-gal is not apparently degraded.

The expression of the *hmg2* gene was analysed by non-radioactive *in situ* hybridisation, that determined the formation of a purple precipitate in correspondence to the transcript (see Materials and Methods). We detected very high levels of hmg2 transcript in pachytene spermatocytes (filled arrow in Fig.34a). Purple colour was also found in round spermatids (not shown). The counterstain with DAPI shows clearly that spermatogonia and elongating spermatids do not contain transcript (arrows in Fig.34b).

2.16 Electron microscope analysis of seminiferous tubules in -/- mice

Since severe defects in cellular organisation were found by histological analysis, seminiferous tubules of -/- mice were analysed by electron microscopy (EM).

EM images revealed that in several spermatids the acrosome is completely detached from the nucleus, as shown in Fig.35. The acrosome is a Golgi-derived structure containing enzymatic granules that play a crucial role during egg fertilisation. Usually, during acrosome formation, the Golgi apparatus and the cytoplasm migrate to opposite sides of the nucleus. At the same time, the acrosome elongates and originates a compact structure whose particular shape appears to be crucial for the mobility of spermatozoa and for the fertilisation of eggs. Fig.35 shows a late spermatid of *hmg2*$^{-/-}$ mice: the nucleus appears normally elongated and contains highly compact chromatin, while the acrosome is completely detached from the nucleus; the space between the acrosome and the nucleus contains unidentified, vesicular structures (asterisks in Fig.35).

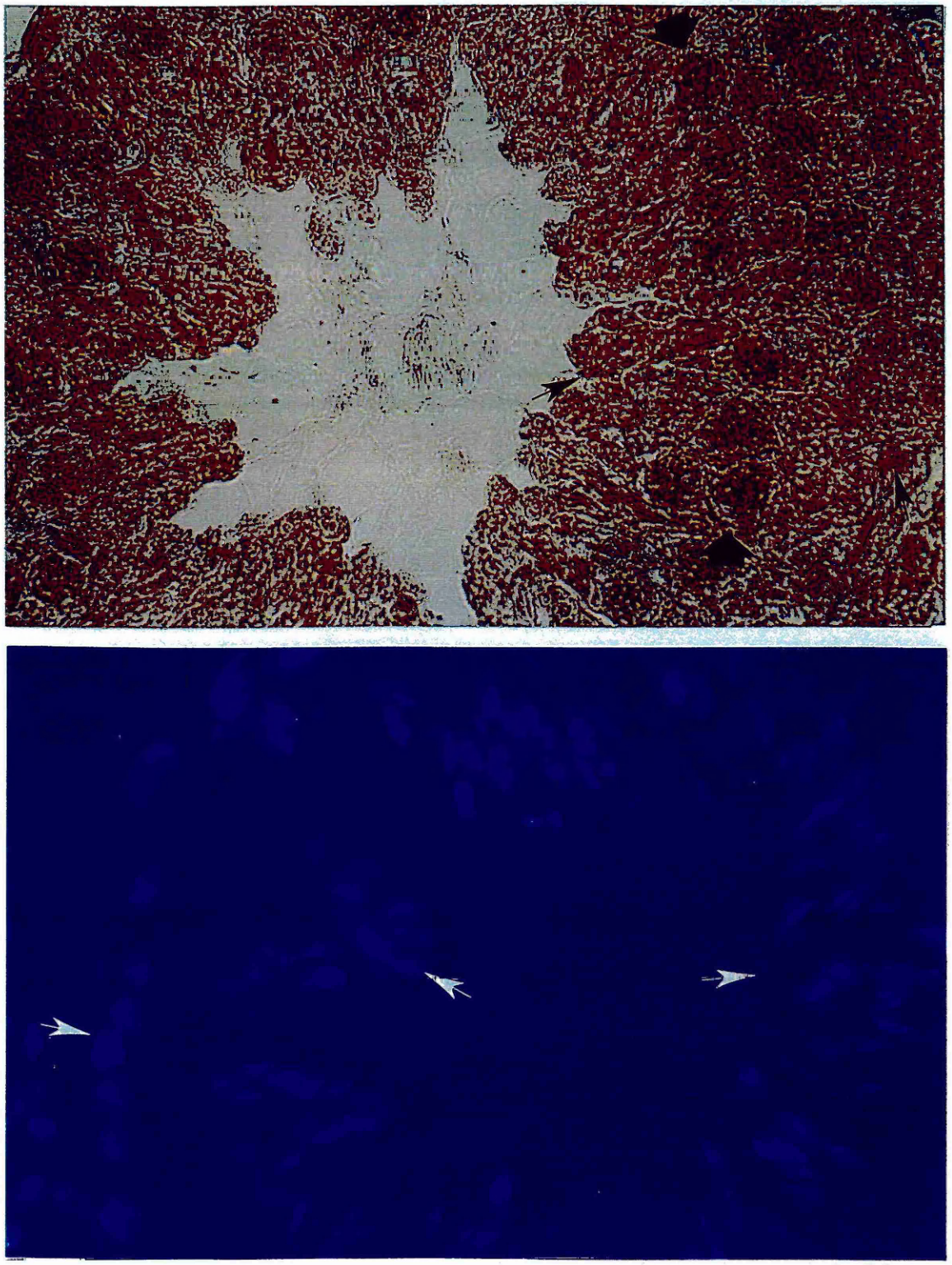


Fig.34 (a) In *situ* hybridisation performed on testis sections from 60 days old wild type mouse. Intense purple color is present in pachytene spermatocytes (filled arrow). Signal is absent in spermatogonia and elongating spermatids. (b) The same section was counter stained with DAPI, unfortunately the section was photographed slightly shifted to the left. Arrows indicate the cells that don't contain hybridisation signal: spermatogonia and elongating spermatids

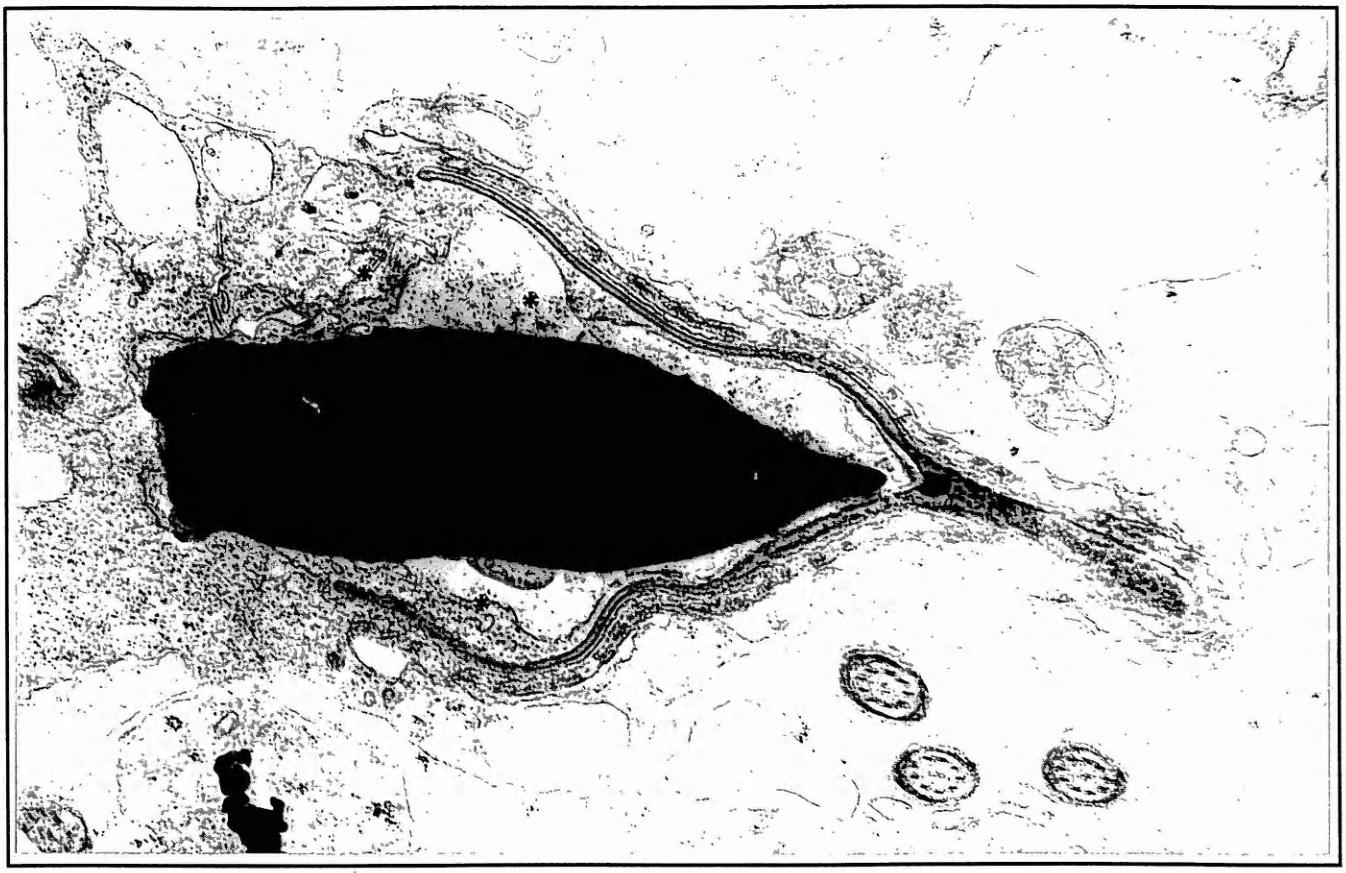


Fig.35 Electron microscopy of an elongating spermatid. The nucleus appears condensed. The arrow indicates that the acrosome is completely detached from the nucleus. The space between the acrosome and the nucleus contains vesicles and empty organelles (asterisks).

Moreover, in certain late spermatids we found one or more flagellar midpieces embedded inside the cytoplasm (not shown).

EM analysis revealed also the presence of big gaps between spermatids and Sertoli cells, as shown in Fig.36. Sertoli cells can be distinguished from the other cellular component of the tubules by the presence of several enlarged mitochondria and by a highly developed smooth endoplasmic reticulum (asterisks in Fig.36). EM images revealed also that Sertoli cells contain an abnormal smooth endoplasmic reticulum, a condition generally associated to suffering cells, as well as the presence of big intercellular gaps, that in general are present when the cells are dying. Usually, wild type cells of the seminiferous tubules contain so-called "multivesicular bodies", whose function is still unknown. In -/- Sertoli cells the multivesicular bodies appeared less regularly shaped than in wild type and so enlarged that they form big vacuoles.

2.17 -/- spermatozoa have reduced motility but are competent to fertilise eggs

As described before, in completely sterile knock-out mice, about 15% of seminiferous tubules are completely degenerated, while the others are apparently normal. In fact, although most of them contain an increased number of apoptotic cells, spermatozoa are produced. Therefore, we decided to investigate why spermatozoa from completely sterile mice are not able to fertilise eggs. Spermatozoa were collected either from the epididymis or by flushing the uterus of females mated with -/- and +/+ males. They were analysed by phase microscopy: in most samples from mutant mice a significant percentage of spermatozoa were not motile. On the contrary, spermatozoa from wild type mice were always motile. In general the shape of -/- spermatozoa is different from that of wild type: tails are shorter and the hook shaped head is not so sharp (not shown).

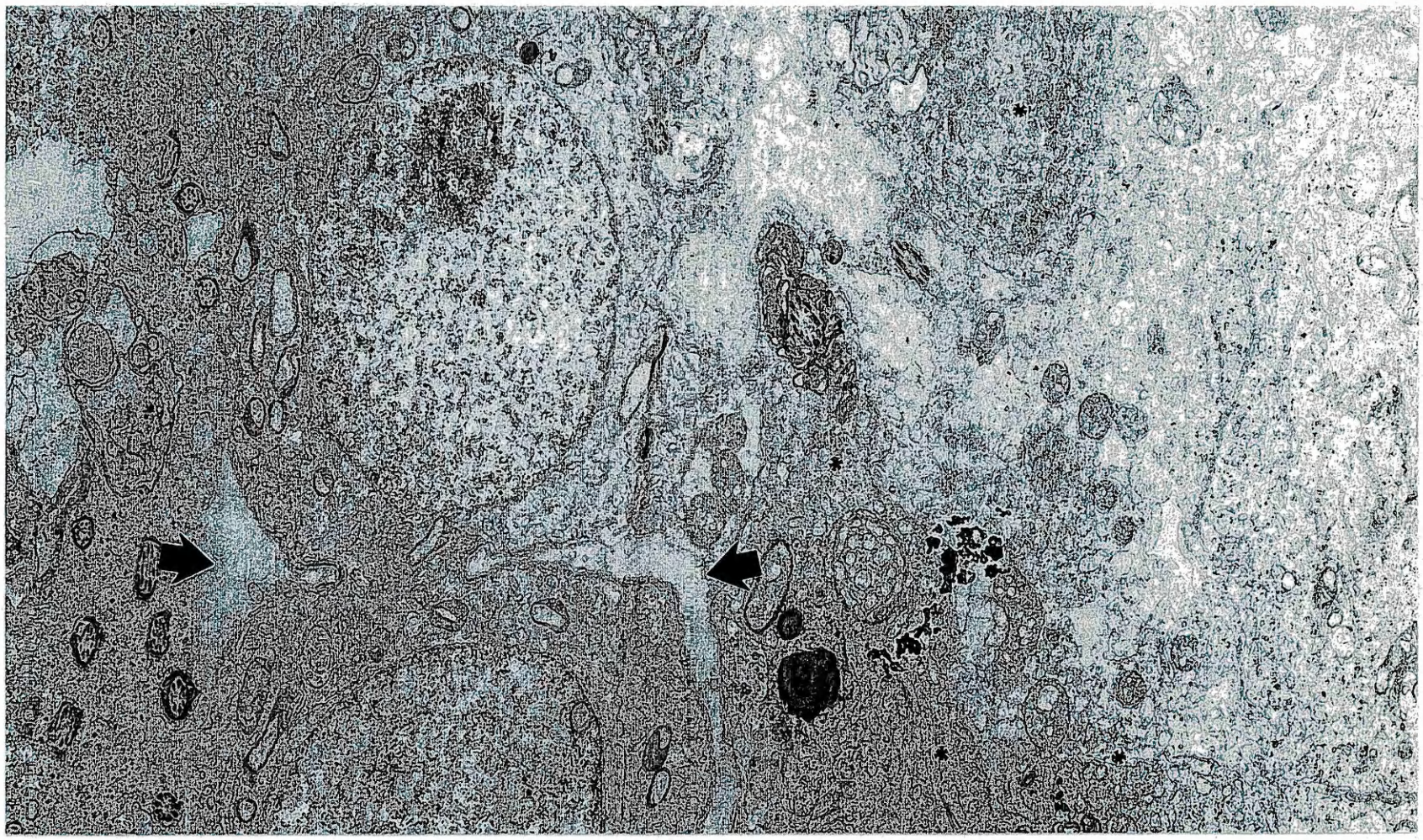


Fig.36 Electron microscopy of a Sertoli cell surrounding two spermatids. Arrows indicate the big gaps between Sertoli cell and spermatids. Sertoli cell contains highly developed smooth endoplasmic reticulum (asterisks). In general the presence of big gaps between cells indicate that a cell is suffering.

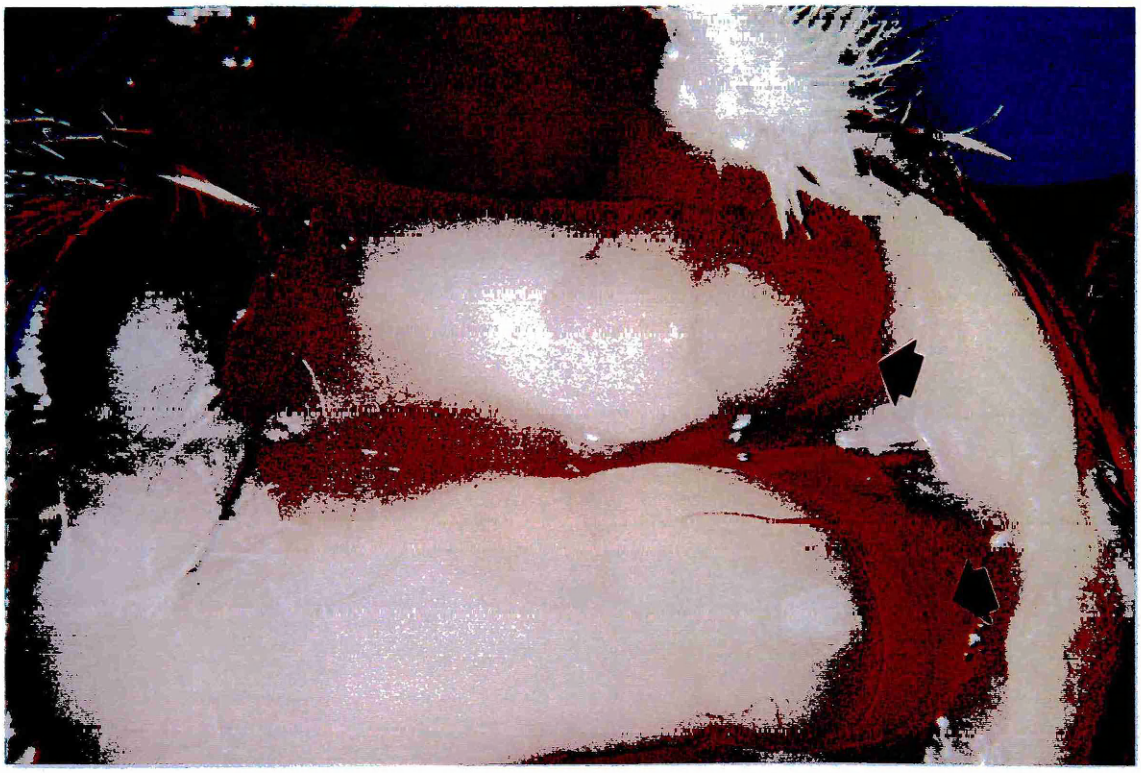
Spermatozoa isolated from -/- and +/+ mice were used for *in vitro* fertilisation assays. 10^6 spermatozoa were added to oocytes collected from the oviduct of superovulated females and were incubated for 6 hours in a humidified 37°C incubator with 5% CO_2. Fertilised eggs were distinguished by the presence of male pronuclei and two polar bodies. Spermatozoa from different knock-outs are able to fertilise eggs with the same efficacy than wild type: in both cases about 50% of the eggs were fertilised.

2.18 +/+ and -/- embryonic fibroblasts respond in a similar way to androgens

In various HMG2 deficient mice, testes were in the inguinal ring instead of the correct scrotal position (Fig.37). As described in the discussion, some authors suggest that the descent of testes from the inguinal ring to the scrotum is regulated by androgens; in fact, mice with abnormal androgen receptor (AR) function have testes in the inguinal ring (Hutson, 1986; Hutson and Donahoe, 1986; Perez-Palacios and Jaffe, 1972). However, mice with no AR also have other defects as the absence of epididymis, vas deference and coagulating glands. *hmg2*$^{-/-}$ mice have none of the above listed morphologies.

Several findings suggest that our knock-out mice have hormonal defects. In fact, the histology of the testis is very similar to that of transgenic mice overexpressing ABP protein and of mice lacking of the estrogen receptor (ER) that regulates the availability of androgens. In both ER and *hmg2* knock-outs the number of dysmorphogenic seminiferous tubules increases with the age of the animal.

It was previously demonstrated that HMG2 interacts with the androgen receptor (AR) and it was proposed that HMG2 acts as co-factor in the transcriptional activation mediated by AR (Boonyaratanakornkit et al., 1998). Therefore, we decided to test AR activity in embryonic fibroblasts isolated from wild type and knock-out mice, by transient co-transfection assay.



a

b

Fig.37a Position of testes in a wild type mouse. Arrows indicate that the testes are in the scrotum. Fig.37b Position of testes in a knock-out mouse. Testes are not in the scrotum. Arrows indicate the putative position of testes.

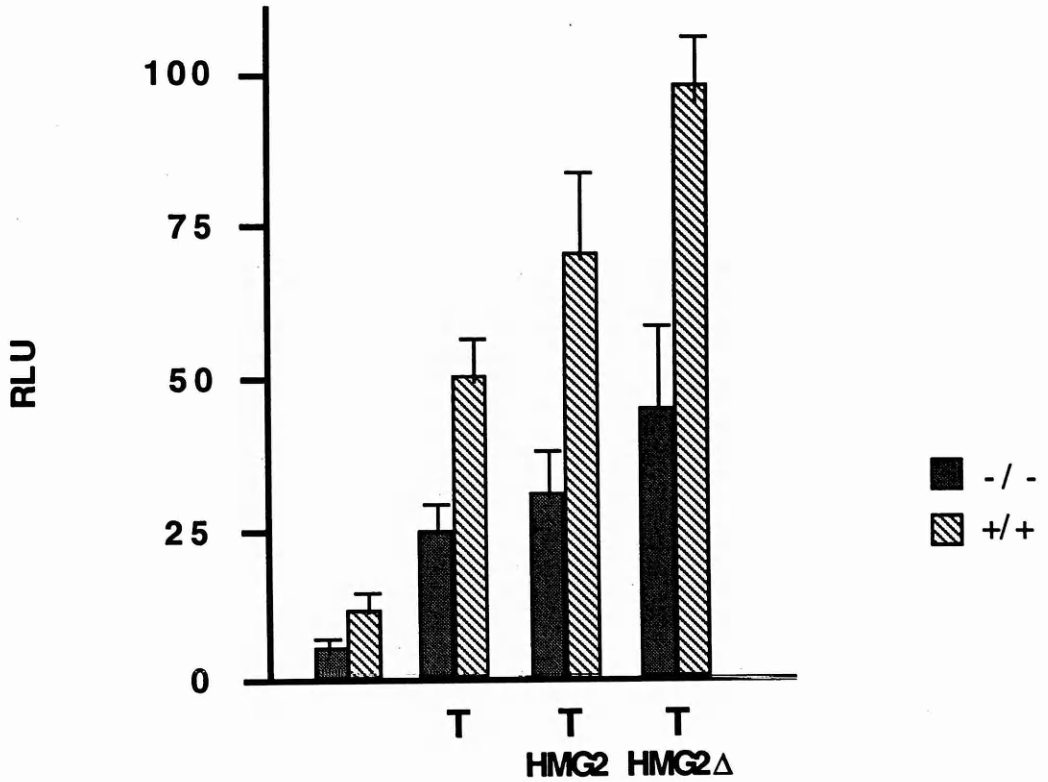


Fig.38 Transient transfection of MEFs with 3μg of a plasmid containing the mouse androgen receptor (mAR) and 6 μg of a reporter plasmid containing the luciferase gene under the control of progesterone response element (PRE-TK2-LUC). Cells were induced with testosterone. Both wild type and knock-out mice response to induction. Co-transfection of pHMG2 determines a modest increase of the trascritional activation in both -/- and +/+ MEFs. Co-transfection of a truncated form of HMG2 determines a higher increase of the transcriptional activation.

Mouse Embryonic Fibroblasts (MEFs) were isolated from wild type and knock-out 16 days old embryos. Primary cultures were co-transfected with 6 μg of PRE-TK2-LUC, a plasmid carrying the luciferase gene under the control of the progesterone response element (PRE) and with 3 μg of mAR, a construct expressing the mouse androgen receptor gene. MEFs were grown in high glucose DMEM medium supplemented with 10% of hormone depleted serum, and 48 hours after transfection were induced with 10^{-7} M testosterone for 24 hours. The result of transfections is shown in Fig.38. In all tested conditions -/- cells show reporter transcription that is always half of the level measured in +/+ cells. Testosterone induction determines an increase transcription of 4-fold in both +/+ and -/- MEFs in comparison to non induced cells. MEFs were then co-transfected with a construct containing the whole *hmg2* gene under the control of its own promoter, pHMG2. Transient expression of HMG2 in both +/+ and -/- fibroblasts determined a modest increase of the transcriptional activation mediated by AR. We further co-transfected cells with pHMG2Δ, that expresses a truncated form of HMG2 lacking the acidic tail: both +/+ and -/- MEFs respond to testosterone induction better than MEFs transfected with pHMG2.

No major differences between +/+ and -/- MEFs were found in the response to androgen induction. This result apparently suggests that in the absence of HMG2 the androgen receptor performs correctly its function possibly because HMG1 is present. Nevertheless, we cannot exclude the possibility that most defects found in -/- mice are due to hormonal problems. In fact, embryonic fibroblast may not be the most appropriate to observe AR-dependent effects. Since during spermatogenesis *hmg2* is highly and specifically expressed in germ cells, they could be the best cellular type to use to test differences in the response to testosterone. Unfortunately, this kind of cells are unavailable for transfection assays.

3. DISCUSSION

The main purpose of my thesis work was to elucidate the role of HMG2 protein in the cell and to identify possible differences between HMG1 and HMG2. Therefore, the mouse *hmg2* gene was cloned and characterised, and functional assays were done.

I first documented that HMG2 is functionally equivalent to HMG1. Tiziana Bonaldi in our laboratory showed that both HMG1 and HMG2 proteins interact with HOX proteins. I further demonstrated that HMG2 is similar to HMG1 in the enhancement of transcriptional activation mediated by HOXD9: co-transfection of HMG2 and HOXD9 determines a 6-fold increase in the transcriptional activation of a reporter controlled by HOXD9 element. Not even differences in the length of the acidic tails of HMG1 and 2 determine functional differences. Moreover, in collaboration with Dean Edwards's group, we demonstrated that HMG2 is a co-activator of transcription mediated by steroid hormone receptors, just like HMG1 (Boonyaratanakornkit et al., 1998). I then analysed the association of HMG2 with chromatin. I showed that in contrast to the behaviour of other chromatin proteins, as histone H1 and HMG-I(Y), HMG2 is not stably associated to interphase or metaphase chromosomes; the same result was obtained for HMG1 by Falciola et al. (1997). Thus, in all biochemical and cellular assays, HMG2 and HMG1 appear to play the same function. Data previously published are also in agreement with my results: both HMG1 and HMG2 proteins are able to recognise cruciform DNA, can replace bacterial HU protein in site specific recombination, and are involved in V(D)J recombination (Aidinis et al., 1999; Bianchi, 1988; Schatz et al., 1989).

A unique and substantial difference between HMG1 and HMG2 was found: HMG1 is ubiquitously expressed, while HMG2 is differentially expressed. In adult liver and brain, neither hmg2 transcript nor HMG2 protein are detectable. Low levels of

protein and messenger were found in other adult tissues; high amounts are only found in thymus, spleen and testis. These tissues are sites of high cellular proliferation. Moreover, actively growing NIH3T3 cells express higher level of *hmg2* than growth-arrested cells. These data are compatible with the hypothesis that *hmg2* expression correlates with cell proliferation. However, *in situ* hybridisation analysis of adult testis revealed that in this organ high expression of *hmg2* is not associated with cell proliferation. In fact, in the only proliferating cellular component of the testis, the spermatogonia, *hmg2* is not expressed or its expression is so low that is not detectable by this technique. On the contrary, *hmg2* is highly expressed in pachytene spermatocytes and early spermatids, the cells undergoing meiosis (Fig.39). Moreover, in the CNS *hmg2* is highly expressed in both proliferating and differentiated neurons. Our data run contrary to the notion that HMG2 is required in actively dividing cells (Seyedin and Kistler, 1979; Shirakawa and Yoshida, 1995; Yamazaki et al., 1995).

We have little understanding of the mechanisms that differentially control the gene expression of *hmg2* and *hmg1*. For example, Fig.39 shows clearly that during spermatogenesis *hmg2* and *hmg1* have expression patterns completely different: contrary to HMG2, HMG1 protein is present at low and constant level in all cell types with the exception of elongating spermatids and spermatozoa.Post-translational control could play a large role: sequence comparison between human and mouse HMG2 revealed that 3'-UTRs are more than 90% of identical. In fact, all 3'-UTRs of HMG2 cDNAs cloned from different species are similar (Davis and Burch, 1992), and are totally different from 3'-UTRs of HMG1s (that also form a tightly conserved ensemble). Future experiments will address this point.

The absence of HMG2 does not determine dramatic phenotypes, in contrast to the absence of HMG1. Both *hmg2*+/- and -/- mice develop normally and have no obvious problems. This result conclusively disproves that HMG2 is absolutely required for cell

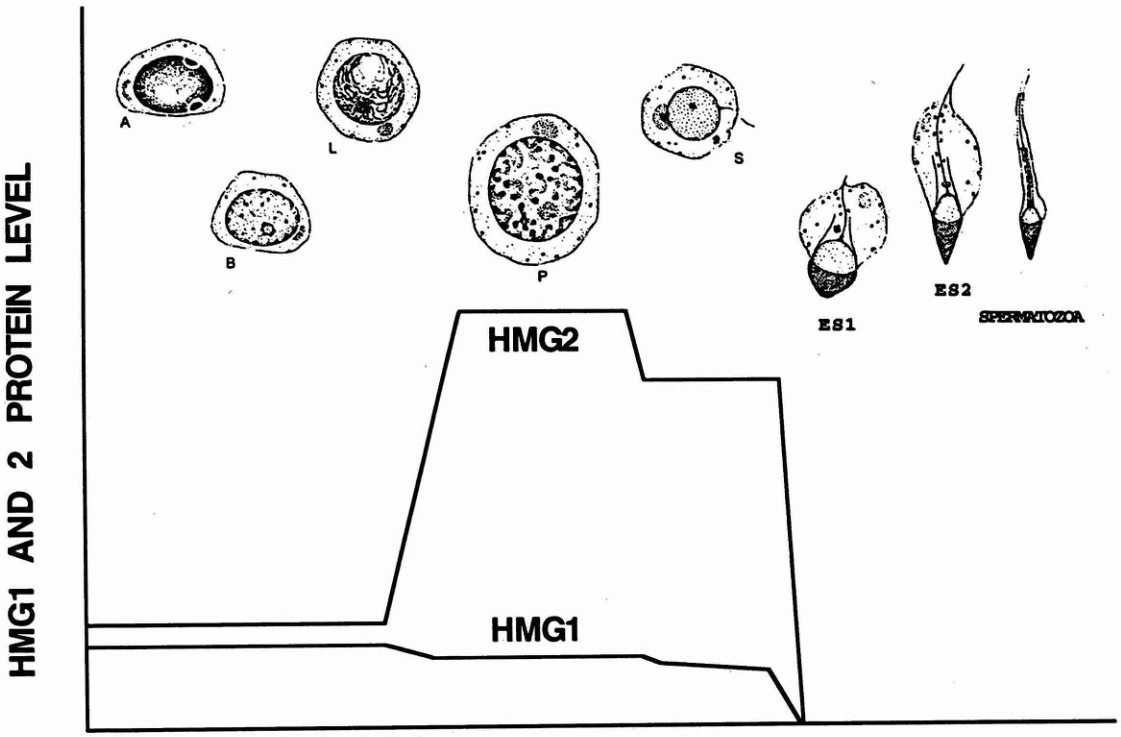


Fig.39 Distribution of HMG2 and HMG1 during spermatogenesis. High level of HMG2 were found, by immunohistochemistry analysis, in pachytene spermatocytes and round spermatids. No protein was found in elongating spermatids and in spermatozoa. HMG1 level is constant. The expression profile of HMG1 shown in this grafic was described in Bucci et al. (1984). A and B, spermatogonia of type A and B. L, leptotene spermatocytes. P, pachytene spermatocytes. S, round spermatids. ES1 and ES2, elongating spermatids, two different stages of maturation.

proliferation, as suggested by Seyedin et al. (1979) and by Shirakawa et al. (1995). *hmg2* mutant mice were analysed for immunological and homeotic abnormalities because *in vitro* evidences indicated an interaction of HMG2 with RAG1 and HOX proteins, but no defects were found. *hmg2* mutant mice only show a partially reduced male fertility. This defect is of variable penetrance, because some mice are totally sterile, but some are normally fertile. The histology of testes from -/- mice revealed an age-related degeneration of part of the seminiferous tubules, and increased apoptosis in apparently normal tubules. Spermatozoa are produced, but some are morphologically degenerate and most are immobile. Nonetheless, there are enough functional spermatozoa to fertilise oocytes *in vitro*.

The testis is one of the few organs where HMG2 is present in the adult. There is a wave of HMG2 synthesis starting in primary spermatocytes, while elongated spermatids contain no HMG2. This tight regulation of expression, and the high level of HMG2 attained, suggests that HMG2 plays a specific role in male cells undergoing meiosis. Mature germ cells are produced in *hmg2*-/- mice, so that a direct role in recombination and chromatid segregation appears unlikely. Most likely, HMG2 facilitates the access of interacting transcription factors to genes that are actively transcribed during the progress of meiosis. Again, the effect is probably moderate, because in *hmg2*-/- mice some spermatozoa are correctly differentiated and functional. Spermatocytes are not the only cells affected by the absence of HMG2: some Sertoli cells show vast vacuoles and lacunae. This might be due directly to the absence of HMG2 in knock-out Sertoli cells, but could also derive from defective interactions with differentiating spermatocytes.

Defects similar to these were found in testes of mice lacking the estrogen receptor (Eddy et al., 1996), and of mice overexpressing the rat androgen binding protein (ABP) (Hammond, 1990). Estrogens regulate androgen production (Eddy et al., 1996), while ABP is a glycoprotein secreted by Sertoli cells, that binds with high affinity androgens produced by Leydig cells, and is proposed to regulate spermatogenesis and sperm

maturation by maintaining high androgens levels in the testis and epididymis (Anthony et al., 1984). Moreover, in several *hmg2*-/- mice we found testes in the inguinal ring instead of the correct scrotal position (Fig.37). Testicular descent, from the abdomen to the inguinal ring is proposed to be under the control of the insulin-like hormone (Insl3) (Nef and Parada, 1999), whereas the final descent from the internal inguinal ring to the scrotum is androgen-dependent: in male mice with testicular feminisation, the testes are located lateral to the bladder neck (Hutson, 1986; Hutson and Donahoe, 1986; Perez-Palacios and Jaffe, 1972).

The previous observations, and the enhancement brought about by HMG2 on the transactivating activity of the androgen receptor (AR) (Boonyaratanakornkit et al., 1998), suggest that AR function is impaired in the testes of *hmg2*-/- mice. However, in transfection assays we found only two-fold differences in the activity of the AR in *hmg2* +/+ and -/- mouse embryonic fibroblasts (MEFs). We do not consider this difference to be significant. Possibly, MEFs do not reproduce the intracellular milieu of the real target cells in the testis, and actually MEFs contain a much higher amount of HMG1 than HMG2.

In conclusion, my work suggests that HMG1 and HMG2 can vicariate each other in most cells, and that the total amount of HMG1 and HMG2 is relevant. In fact, the main defects found in *hmg1* -/- knock-out mice are tissues where HMG2 is not present. For example, *hmg1* -/- mice mutant newborns die because of hypoglycaemia as they are unable to metabolise glycogen accumulated in the liver, where HMG2 is absent. Conversely, *hmg2* -/- mice have defects in the testes, where HMG2 is more abundant than HMG1. The analysis of double knock-out mice we will obtain by crossing *hmg1* +/- and *hmg2* +/- will test our hypothesis.

4 MATERIALS AND METHODS

4.1 *E. coli* strains

DH5α: Strain with genotype *supE44 DlacU169 (F80 lacZDM15) hsdR17 recA1 end A1 gyrA96 thi-1 relA1.* A recombinant-deficient suppressing strain used for plating and growth of plasmids.

XL1-Blue: Strain with genotype *supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac⁻ F'(proAB⁺ lacIᵠ lacZDM15* Tn*10(tetʳ)).* A recombination-deficient strain that will support the growth of vectors carrying some amber mutations. Transfected DNA is modified but not restricted. The F' in this strain allows blue/white screening on X-GAL and permits bacteriophage M13 superinfection.

HB101: *supE44 hsdS20($r_B^-m_B^-$) recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1.* A suppressing strain commonly used for large-scale production of plasmids. It is an *E. coli* K12 X *E. coli* B hybrid that is highly transformable.

JM110: *dam⁻ dcm⁻supE44 endA1 hsdR17 thi leu rpsL lacY galK galT ara tonA thr tsx D(lac-proAB) F'(traD36 proAB⁺ lacIᵠ lacZDM15).* A strain that will not modify *Bcl*I sites and will support growth of vectors carrying amber mutations.

4.2 Liquid media

To growth *E. coli* we used LB (Luria-Bertani) Medium that when necessary was supplemented with ampicillin at a concentration of 100mg/ml as selective agent.

LB (per litre):

bacto-tryptone	10g
bacto-yeast extract	5g
NaCl	10g

4.3 Protocol of plasmid DNA extraction

Plasmid DNA was prepared using two classical techniques of plasmid extraction. The first was used to prepare plasmid DNA from small amounts of many different cultures (1 to 24) of plasmid-containing bacteria miniprep and the second one was used to isolate large quantity of plasmid DNA. In both cases the alkaline lysis procedure was used (Maniatis et al., 1989). For the production of large amounts of DNA the purification was done using anionic exchange cartridges produced by QIAGEN.

4.4 Oligonucleotides used for the amplification of introns 3 and 4

Sequence of the oligonucleotides used for the amplification of intron 3 and 4 of the mouse *hmg2* gene:

INT3for (coding strand)	5'-CCGAAGAGACCACC**G**-3'
INT3rev (non-coding strand)	5'-GGAAGAAGGCAGAC**CT**-3'
INT4for (coding strand)	5'GGAGAAGTATGAAAAG**GT**-3'
INT4rev (non-coding strand)	5'GGTATGCAGCAATATC**C**-3'

The oligonucleotides were provided by Genset (France), they are based on the coding sequence of the mouse HMG2 cDNA. They contain at their 3' terminus one or two extra bases corresponding to the canonical splicing junction 5'-GT and 3'-CT (letters in bold).

We performed PCR on mouse genomic DNA from 129SV strain with the following conditions:

DNA	20 ng
INT3for	30 pmol
INT3rev	30 pmol
dNTPs	0.2 mM

DNA polymerase bf (10X)	3 µl
DNA polymerase (Finnzyme)	0.6 units
H_2O	to 30 µl of reaction mixture

PCR conditions: thirty-five cycles of denaturation (30 s at 94°C), annealing (45 s at 59°C for INT3 and 45 s at 56°C for INT4), and extension (30 s at 72°C) on a Hybaid OmniGene machine.

The same reaction was done to amplify the intron 4.

4.5 Isolation of the genomic clones containing the *hmg2* gene

$1x10^6$ phage plaques (corresponding to 4 equivalent genomes) from a mouse genomic library cloned in λ-FIXII provided by Stratagene were screened using intron 4 as a probe. The host cells (XL1-blue) were grown at a OD=0.5 at 600nm; 600µl of these cells were infected with 10^5 phages and incubated at 37°C for 15 minutes. Infected cells are added to 7 ml of top agar (5 g NaCL, 2 g $MgSO_4x7H_2O$, 5 g of yeast extract, 10 g casein hydrolysate and agarose to 0.7%) and plated on petri dishes containing agar. The plates were then incubated for 6-8 hours at 37°C.

DNA was transferred from plaques onto a nylon membrane (GeneScreen *Plus*, Du Pont). The DNA on the membrane was denatured in a 0.5 M NaOH solution and then neutralised in Tris-HCl 0.5 M pH 7.6 plus NaCl 1.5 M. The filters were washed in 2XSSC (1.5 M NaCl and 0.15 M Na citrate) and dried.

The filters were then hybridised with intron 4 as probe. Intron 4 was radioactively labelled using the nonamer primers kit from Amersham and attending the instructions of the Company. The hybridisation was done at 42°C O/N in a mix containing 50% formamide, 5X SSC, 5X Denhard's, and 150µg/ml of herring sperm. The filters were washed twice at 65°C with 0.2X SSC and 0.2% SDS for 45 minutes and then were exposed to X-ray film.

DNA fragments from positive clones were subcloned into pBlueScriptII KS(+) and sequenced with the canonical primers T7 and T3, using in part the sequencing service of PRIMM and in part the Cycle Sequencing Kit (New England Biolabs).

4.6. RNase protection assay and Northern blot analysis

Total cellular RNA was extracted from tissues or cells by the guanidine isothiocyanate technique as described by Chomczynski and Sacchi (1987). For RNase protection assays we used kit from Ambion. 10 µg of total RNA from embryonic fibroblasts was incubated overnight at 45°C with $1.5x10^5$ cpm of riboprobe obtained by *in vitro* transcription of a 270-bp SmaI-XbaI fragment cloned into the pBlueScript II KS(+) plasmid, which spanned a region 250 bp upstream and 20 bp downstream of the exon 1 of the *hmg2* gene.

Transcription bf (10X) (Boehringer)	2 µl
4 mM ATP, CTP, GTP	2 µl
UTP (10 µM)	2.5 µl
RNase inhibitor (Promega)	20 units
DNA	1.5 µg
T3 or T7 RNA pol (Boehringer)	20 units
H_2O	to 20 µl

After RNase treatment the protected fragments were analysed on a 6.5% sequencing gel.

For RNA blot analysis, 5µg of total RNA isolated from different adult tissues and embryonic stages were separated on a 1.2% agarose/formaldehyde gel, transferred onto a GeneScreen Plus membrane (Du Pont) and probed with ^{32}P-labelled HMG2 cDNA.

4.7 Mapping of the mouse *hmg2* gene

We found a polymorphism in the length of the intron 4 between recombinant inbred strains, that allows us to map the mouse *hmg2* gene. The DNA was provided by The Jackson Laboratory and was prepared from the backcross progeny of (C57BL/6J X *M. spretus*) F_1 hybrid female mice mated with C57BL/6J males, designated BSS. We performed a radioactive PCR to amplify the intron 4 on the progeny BSS DNA: 50µl reaction mixtures contained 125ng of mouse genomic DNA, 5µl of 10X Promega DNA polymerase buffer, 1.5mM $MgCl_2$, 0.2mM dNTPs, 1µM INT4for, 1µM INT4rev, 1µCi of [^{32}P]dCTP. 2.5 units of AmpliTaq DNA polymerase (Promega) were added after an hot start at 94°C for 5' and at 80°C for additional 3' and the PCR was done as follows: 94°C denaturation reaction for 30 s, 62°C annealing reaction for 30 s and 72°C extension reaction for 30 s (2 cycles). Other 2 rounds of 2 cycles each were performed at 60°C and 58°C annealing reaction, respectively, followed by 24 cycles at 56°C of reaction annealing with a final extension reaction at 72°C for 4 minutes after the last cycle. The PCR products were separated on a 6% nondenaturing polyacrylamide gel containing 1X TBE. The gel was then dried and exposed to X-ray film.

4.8 Differential permeabilisation of cells

NIH3T3 fibroblasts were grown in 75-cm^2 flasks to 70 % confluence and then were exposed to 50 ng/ml nocodazole for 16 hours. Cells blocked in mitosis were detached by manual shaking of the culture flasks and recovered by centrifugation. Non-mitotic cells were recovered by mild trypsinisation and centrifugation. The two cell populations were checked for the presence of condensed chromosomes by staining of small aliquots with Hoechst 33258: more than 95% of the cells detached by shaking were found to have condensed chromosomes, against less than 2% of the cells that resisted detachment. Metaphase and interphase cells were permeabilised by adding ice-cold

transport buffer (TB buffer, 20mM Hepes, pH7.3, 110mM K-acetate, 5mM Na-acetate, 2mM Mg-acetate, 1mM EGTA, 2mM DTT, and pepstatin, antipain, and leupeptin at 1μg/ml each) containing 0.2% NP-40. Cells were then centrifuged and the supernatants were recovered while the pellets were resuspended in an equal volume of ice-cold TB buffer containing 0.1% NP-40, 10mM $MnCl_2$, and 20μg/ml DNaseI. The suspensions were incubated 10 minutes at 37°C and analysed by Western blotting.

4.9 Western blotting

Cell lysates were prepared from different adult tissues and from different cell lines. Prior to lysis, frozen tissues were powdered by mechanic pressure and then were pottered in an adequate volume of lysis buffer determined in function of the weight of the tissue. The lysis of cells growing in a monolayer was done directly by adding lysis buffer to the petri dish. In both the cases lysis buffer contained 10mM Hepes, pH 7.3, 0.1mM DTT, 1mM EDTA, 100mM NaCl, 0.1% Triton X-100, 0.2 mM PMSF and pepstatin, antipain, and leupeptin at 1μg/ml each. SDS-PAGE loading buffer was added to cell extracts and the samples were loaded on a Tricine-SDS-PAGE. The separating gel consisted of polymerised 10% acrylammide-bisacrylammide (ratio 29:1) in 1 M Tris pH 8.45, 0.1% SDS. The stacking gel consisted of 3.6% acrylammide-bisacrylammide, 0.75 M Tris pH 8.45, 0.1% SDS. The anode buffer contained 0.2 M Tris pH 8.9, the cathode buffer contained 0.1 M Tris, 0.1 M Tricine, and 0.1% SDS final pH8.25. The gels were run at 5-15 Volts/cm.

The proteins separated in the SDS-PAGE were transferred on a nylon membrane (Millipore) in a transfer apparatus in transfer buffer (25mM Tris, 192mM glycine, 20% methanol). Filters were then incubated in blocking buffer (5% dry non-fat milk in TBST, 20mM Tris-HCl pH 7.4, 136mM NaCl, and 0.1% Tween20). The filter was then incubated with the chicken anti-HMG2 antibody at a final concentration 1-2μg/ml diluted in TBST containing 0.1% BSA, washed and incubated with a rabbit anti-chicken

antibody conjugated with Alkaline Phosphatase, which was revealed with CDP-Star (Boerhinger Mannheim) for chemioluminescence.

4.10 Constructs and plasmids

pHMG2 is a construct that contains the whole *hmg2* gene under the control of its own promoter. A *Eco*RI 4.8 kb long fragment isolated from 129SV mouse genomic library (λ-FIXII) was cloned into the pBS-KS(+). This fragment contains a genomic region 1.2 kb long upstream the exon 1 of *hmg2,* the whole gene and, over the 3'-UTR, a genomic region 1 kb long.

p5kbXbaI is a construct that contains the whole *hmg2* gene. A *Xba*I 5 kb long fragment was cloned in the pBS-KS(+). It contains 270 bp upstream the exon 1, the gene and a 2.5 kb long region downstream the 3'-UTR of the gene.

pHMG2Δ and pHMG2sw are two constructs that derive from pHMG2. The first carries a mutation that introduces a stop codon at the beginning of the acidic tail of HMG2. The mutagenesis was done by PCR. Two pairs of complementary oligonucleotides were synthesised on the sequence of exon five in order to mutagenise glu 186 into a stop codon (GAA to TAA).

Sequence of the mutated primers, the bases mutated are in bold:

Zdirhmg2	5'-GAAGAAGAAC**TAA**CCAGAAG-3'	(coding strand)
Zrevhmg2	5'-CTTCTGG**TTA**GTTCTTCTTC-3'.	(non-coding strand)

Zrevhmg2 was used to amplify a 94 bp long fragment in couple with StyIdirhmg2 5'-GGATATTGCTGCATACCG-3' synthesised upstream on the coding strand. Zdirhmg2 was used to amplify a 296 bp long fragment in couple with HindIIIrevhmg2 5'-ATCAAAAGCTTTAAATCTAAC-3' synthesised downstream on the non-coding strand. To amplify the two fragments the same PCR conditions were used: thirty-five cycles of denaturation (30 s at 94°C), annealing (45 s at 54°C) and extension (30 s at 72°C). An equal quantity of PCR products was mixed and used to amplify a 371 bp long

fragment using the external primers StyIdirhmg2 and HindIIIrevhmg2. For the second amplification the following PCR conditions were used: three cycles of denaturation (30 s at 94°C), annealing (45 s at 50°C) and extension (30 s at 72°C), followed by thirty-five cycles of denaturation (30 s at 94°C), annealing (45 s at 54°C), and extension (30 s at 72°C). The PCR reaction was done using the high fidelity Deep*Vent* DNA polymerase from New England Biolabs, in order to avoid the introduction of mutations during the amplification. The PCR product was then digested with *Sty*I and *Hind*III restriction enzymes and inserted into the pHMG2 previously digested with the same enzymes. The pHMG2Δ was sequenced in order to verify the presence of the mutation in the correct position and at the same time to verify the absence of random mutations eventually introduced during the amplification.

pHMG2sw derives from pHMG2 and carries the acidic tail sequence of the *hmg1* gene instead of the acidic tail sequence of *hmg2*. As for pHMG2Δ, pHMG2 was digested with *Sty*I and *Hind*III eliminating the acidic tail and the 3'-UTR of *hmg2* and a PCR product carrying the acidic tail of *hmg1* and the 3'-UTR of *hmg2* was cloned inside. The acidic tail of *hmg1* was amplified using a couple of mutated oligonucleotides synthesised on the sequence of exon five. The primer on the coding strand (hmg1StyImut) carries a mutation that introduces a restriction site for *Sty*I, while the primer on the non-coding strand (ac.t.hmg1NheImhmg1-rev) carries a mutation that introduces a restriction site for *NheI*, and their sequences are reported below:

hmg1StyImut 5'-CTACAGAGCCAAGGGAAAACCT-3'

ac.t.hmg1NheImhmg1-rev 5'-CTGCGCTAGCACCAACTTA-3'

To amplify the 3'-UTR of *hmg2* two couples of primers were used: the oligonucleotide on the coding strand carries a mutation that introduces the same restriction site introduced on the acidic tail of *hmg1* by hmg1NheImrev (hmg2NheImdir), while the primer on the non-coding strand is the same external primer used to create pHMG2Δ,

HindIIIrevhmg2: 5'-ATCAAAAGCTTTAAATCTAAC-3'

NheImhmg2-dir 5'-AATAAGTGGCTAGCCTAAAGTGTG-3'

HindIIIrevhmg2 5'-ATCAAAAGCTTTAAATCTAAC-3'

The same PCR conditions were used to amplify the acidic tail of *hmg1* and the 3'-UTR of *hmg2*: thirty-five cycles of denaturation (30 s at 94°C), annealing (45 s at 56°C), and extension (30 s at 72°C). The two PCR products, 182 bp and 214 bp long respectively, were digested with *NheI* restriction enzyme. An equal quantity of digested products were mixed together and amplified with the external primers, hmg1StyImut and HindIIIrevhmg2. The resulting PCR product, 376 bp long, was digested with *Sty*I and *Hind*III restriction enzymes and subcloned in the pHMG2 previously digested with the same enzymes.

GEPPO, GEPPINO and GEPPETTO are constructs that carry different deletions of the promoter of the *hmg2* gene. GEPPINO derives from an intermediate construct generated to make the vector for the gene targeting, p5kbXbaI, and it contains the *sh-ble:LacZ* gene put in place of the coding region of HMG2 and fused to the ATG of *hmg2* (see Charpter 4.13). This construct contains 270 bp upstream to exon 1 of *hmg2*. GEPPETTO was constructed in a similar way as GEPPINO, but it derives from pHMG2 and therefore it contains a longer region, 1.2 kb long, upstream of exon 1. GEPPO derives from GEPPETTO and it contains 445 bp upstream of exon 1. The 175 additional bp more were added by PCR. Two oligonucleotides were designed on the sequence of *hmg2*. The primer synthesised on the coding strand carries a mutation that introduces a restriction site for *Xba*I. Sequence of the primers:

XbaIhmg2dirMut 5'-TCGGAGCCATCTAGATCACCTGAAG-3' (coding strand)

XbaIhmg2rev 5'-CGGTTGGAACCGGTCTAG-3' ..(non-coding strand)

The PCR product, that is 211 bp long, was obtained using the following PCR conditions: thirty-five cycles of denaturation (30 s at 94°C), annealing (45 s at 59°C), and extension (30 s at 72°C). The PCR product was then purified and digested with *Xba*I restriction enzyme and cloned into GEPPETTO previously digested with the same

restriction enzyme. The resulting construct was transformed in competent DH5α *E. coli.* DNAs extracted from 4 colonies were sequenced in order to find the construct with the PCR product cloned in the right direction.

The pHMG1 plasmid was constructed by cloning a 12.5 kb *Eco*RI fragment into the pBS-KS(+) vector. The plasmid contains the whole mouse *hmg1* gene under the control of its own promoter.

The pSGD9 construct was gently provided by Dr. Vincenzo Zappavigna, and it contains the human *HOXD9* gene cloned in the SV40 early-promoter-based mammalian expression vector pSG5 (Zappavigna *et al.*, 1991). The pT81HCR reporter construct contains a single copy of the HOXD9 100 bp autoregulatory element HCR, cloned into the polylinker of the pT81luc luciferase reporter vector (Zappavigna et al., 1991). pRSVßgal and pKS are constructs used as internal control of the transfections: the first carries the *LacZ* gene under the control of the LTR of the Rous Sarcoma virus, while the second carries the gene that codes for chloramphenicol acetyl transferase under the control of the SV40 early-promoter.

PRE-TK2-LUC is reporter plasmid that carries the luciferase gene under the control of the progesterone response element gently provided by Dean Edwards (Wagner et al., 1998).

mAR is a plasmid containing the mouse androgen receptor under the control of CMV promoter, cloned into the pcDNAIneo gently provided by Don Tindall (He et al., 1990).

4.11 Cell culture and transfections

The mouse NIH3T3 cell line was maintained in high glucose Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% foetal bovine serum (FBS from GIBCO), 100 IU/ml of penicillin and 100 µg/ml streptomycin, in 5% CO_2 in air

humidified atmosphere. Cells were transiently transfected by calcium phosphate co-precipitation in 9-cm or 6-cm dishes.

In the transfection assay to test the promoter of the *hmg2* gene, 200000 cells were plated in 6 cm dishes and were transfected with 3 μg of reporter plasmid (GEPPETTO, GEPPINO or GEPPO) and with 0.5 μg of pRL-TK as an internal control. Cells were harvested 48-60 h after transfection, lysed and assayed for luciferase and ß-galactosidase expression.

In the HOXD9 transfection assay, 200000 cells were generally transfected with 6 μg of reporter plasmid (pTHCR), 3 μg of expression construct (pHMG1 or pHMG2) and 1 μg of pRSV-ßgal as an internal control.

To test transcriptional activation mediated by the androgen receptor, primary fibroblasts isolated from knock-out and wild type embryos were grown in DMEM, supplemented with 10% FBS and antibiotics. For the transfection 200000 cells were plated in 6 cm dishes and transfected with 6 μg of PRE-TK2-LUC reporter plasmid, 3 μg of mouse AR, 2.5 μg of pHMG2 and 1 μg of pRSV-CAT as internal control. 16 h after transfection, the medium was changed with DMEM supplemented with 10 % FBS deprived of steroid hormones. 48 h after transfection cells were induced with 10^{-7} M testosterone (from SIGMA). The cells were harvested 24 h after induction. Serum deprived of steroid hormones was obtained adding 1 % charcoal and 0.1 % dextran T70 to normal FBS serum. After 10 minutes of incubation the serum was centrifuged for 20 minutes at 4000 rpm.

Generally, all transfections were carried out in triplicate batches, using at least two different DNA preparations of each construct in at least three separate experiments.

4.12 Luciferase, ß-galactosidase and CAT assays

Generally, transiently transfected cells were harvested 48-72 h after transfection. 400 µl of lysis buffer containing 1% triton, 25 mM glycine-glycine pH7.8, 15 mM $MgSO_4$, 4 mM EGTA and 1 mM DTT, were added to cells growing in 6 cm dishes.

For luciferase assays, 20 µl of cell lysates were added to the reaction mix, containing 25 mM glycine-glycine pH 7.8, 2 mM ATP pH 7.8 and 10 mM $MgSO_4$. 100 µl of injection mix, containing 25 mM glycine-glycine and 0.2 mM D-luciferin (from SIGMA) were added to each sample and counted for 10 or 30 seconds in the luminomiter

For ß-galactosidase assay, 130 µl of cell lysates were incubated with an equal volume of the ONPG substrate (o-nitrophenil-ß-D-galactoside, 4 mg/ml) and with 500 µl of Z buffer (0.06M Na_2PO_4, 0.04M NaH_2PO_4, 0.01 M KCl, 1 mM $MgSO_4$ pH 7 and 50 mM ß-mercaptoethanol). The mix was incubated at 30 °C until a clear yellow colour developed. The reaction was stopped adding 300 µl of Na_2CO_3 1M. ß-galactosidase hydrolyses the ONPG substrate generating a yellow compound that adsorbs at 420 nm.

For CAT assay, 80 µl of cell lysates were incubated at 37 °C for one or two hours with 0.5 mM acetyl-coenzyme A and with 3 µCi/ml of D-*threo*-(dichloroacetyl-1-^{14}C) chloramphenicol. After ethyl acetate extraction the samples were dried. The residues were resuspended in ethyl acetate, applied to a silica gel TCL plate and subjected to ascending chromatography. The TCL plate was exposed to X-ray film and the intensity of acetylated forms was measured by densitometry. CAT activity is measured as the quantity of radioactivity incorporated in the mono and diacetyl derivatives.

4.13 The construct for gene targeting: pKO22

hmg2 gene is already expressed in ES cells recovered from four days old blastocysts, as we detected by RT-PCR. We decided therefore to make a replacement vector carrying the positive selectable marker under the control of the *hmg2* promoter.

This kind of construct determines an enrichment for homologous recombinants since the positive selectable gene is expressed only after homologous recombination.

A *Xba*I-*Xba*I 5 kb long fragment containing the whole gene was cloned in pBS-KS+ (p5kbXbaI) and was used to generate the construct for the gene targeting. p5kbXbaI was first digested with *BstE*II and *Stu*I restriction enzymes to eliminate 741 bp long region , from exon 2 to exon 4. With this deletion the first ATG of the *hmg2* gene was maintained in the construct. In place of this region I added a linker that: 1) reproduces the *BstE*II restriction site at 5' and the *Stu*I site at 3', 2) contains at the 5' side a *Nru*I restriction site that will be unique in the final construct, 3) contains at the 3' side a *Not*I restriction site. *Nru*I and *Not*I sites were necessary to have the possibility to insert the selectable gene.

Sequence of the oligonucleotides synthesised to construct the linker

hmg2link 5'GTGACCTCGCGATACTCGAGAAGCTTGCGGCCGCCTGCAGAGG-3'
hmg2knil 5'-CCTCTGCAGGCGGCCGCAAGCTTCTCGAGTATCGCGAG-3'

hmg2link and hmg2knil were annealed and cloned in the p5kbXbaIΔBstEII-StuI. The resulting construct was digested with *Nru*I and *Not*I, and a *Bal*I-*Not*I 3610 bp long fragment containing the selectable gene was inserted. This fragment originated from pUT529Δ, a commercial vector from Cayla, and it contains the *sh-ble* gene that imparts the zeocyn resistance. The selectable gene is fused in frame to the first ATG of *hmg2* and replaces the region that codes for HMG-boxA and part of HMG-boxB. Moreover, the *sh-ble* gene is fused in frame to the *LacZ* gene. Therefore, *LacZ* gene expression is driven by the endogenous promoter of *hmg2*.

The whole insert was digested with *Xba*I and *Xho*I, after mutagenesis of the *Xba*I resticition site in 3', and was cloned into pPNT construct that contains the herpes simplex virus thymidine kinase gene (TK) previously digested with the same restriction enzymes.

This passage was done in order to enrich for homologous recombinants. In the case of random integration of the construct, the TK gene is expressed and the cells are sensitive to acyclovir and its analogs (Fig.40).

The 5' homologous region is 800 bp long and contains 200 bp upstream of the translation start site, exon 1, intron 1 and part of exon 2 until the first ATG of *hmg2*. The 3' homologous region is 3.6 kb long and comprises a genomic region extending from the *Stu*I restriction site in the exon 4 over the 3'-UTR and a long genomic fragment 2.5 kb long beyond the 3'-UTR.

4.14 ES cell culture and production of recombinant clones

R1 Embryonic Stem cells were cultured in DMEM supplemented with 15 % of FBS tested for ES cells (GIBCO), 100 IU/ml of penicillin and 100 µg/ml streptomycin, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 10^{-6} M ß-mercaptoethanol and 10^3 units per ml of leukaemia inhibitory factor (LIF). To avoid differentiation, ES cells were grown not only in the presence of LIF but also onto a layer of embryonic fibroblasts (EMFI cells).

Two rounds of electroporation were made. Generally, in each electroporation assay $3X10^7$ R1 cells were used. Into an electroporation cuvette, $5.6X10^6$ cells resuspended in PBS were incubated with 35 µg of linearised pKO22 and electroporated at 240V, 500µF in a BioRad gene pulser. After electroporation the cells were incubated for 20 minutes on ice and plated in a 10 cm dish. 48 h later, zeocyn and GANC drugs were added at a concentration of 120 µg/ml and 2µM, respectively. Selection was continued for 8 days until non differentiated resistant clones appeared. Colonies were picked with sterile tips, trypsinised and grown on 96-wells plate. After duplication of clones, the master plate was frozen while the duplicate plate was used to extract DNA for the Southern blot analysis.

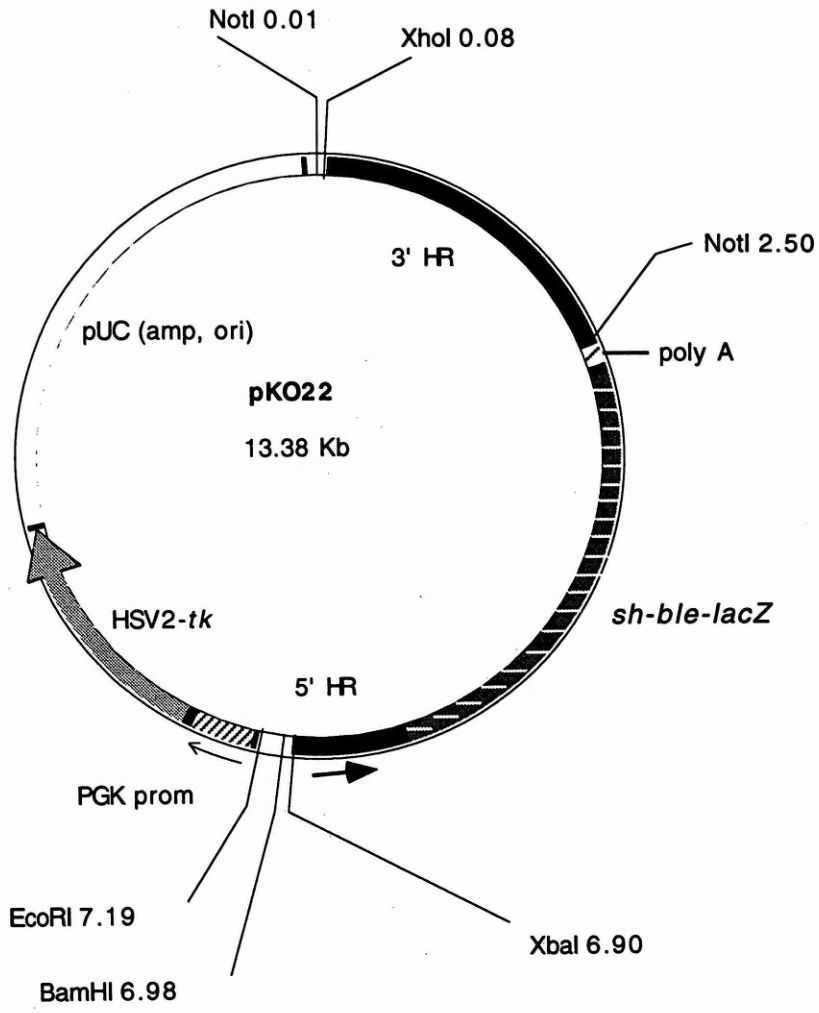


Fig.40 Structure of the construct used to perform the targeting of the *hmg2* gene. Black boxes represent the homologous regions that derive from 5' and 3' of *hmg2* (5' HR and 3' HR). They are 0.8 and 2.5 kb long, respectively. The gray box represents the *sh-ble* gene that imparts the resistence to zeocyn, *lacZ* gene is fused in frame to *sh-ble* gene. Light gray box represents the herpes simplex virus thymidine kinase gene, cells become sensitive to acyclovir when this gene is expressed.

4.15 DNA extraction from ES cell clones and from mouse tails

DNA to characterise ES clones was extracted directly from 96 wells plates. Confluent clones were washed twice with PBS and we added 50 µl of Lysis Buffer (10 mM Tris-HCl pH 7.5, 10 mM EDTA pH 8, 10 mM NaCl, 0.5% sarcosyl, and 1mg/ml proteinase K) per well. The clones were incubated O/N at 60°C in a humidified chamber. The next day, DNAs were precipitated by adding 100 µl of 75 mM NaCl in cold ehanol per well. After a washing in 70% ethanol, the clones were digested directly on the plate adding a digestion mix containing 1X restriction buffer for *Eco*RI, 1 mM spermidine, 100 µg/ml bovine serum albumin, 100 µg/ml of RNAse and 20 units of *Eco*RI from New England Biolabs. The plate was incubated O/N at 37°C in a humidified chamber. Digested DNA was run into a 1% agarose gel and blotted onto a GeneScreen*Plus* membrane. The filter was prehybridised for one hour at 42°C with a mix containing 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulphate, 1 X Denhardt's, 100 µg/ml of denatured herring sperm. A final concentration of probe of 1-4 X 10^5 dpm/ml was added and the hybridisation was continued for 16 hours. The filter was washed twice for 5', with a washing solution containing 2X SSC at room temperature and twice for 20' each in a solution containing 2X SSC and 1% SDS at 65°C. The filter was then exposed to X-ray film.

The probe used to discriminate recombinant clones spans a 700 bp long region in the genome, outside the region of homology at 5' that was used in the construct for gene targeting, but in a region that is comprised in the digested product. This probe recognises two different bands in recombinant clones deriving from the wild type and the mutant alleles, as described in the Results.

Total DNA was also extracted from mouse tails to genotype the animals. Tail biopsies were incubated O/N at 55°C in lysis buffer (50mM Tris-HCl pH 8, 100 mM EDTA, 100 mM NaCl and 1% SDS) with a final concentration of 1 mg/ml of proteinase K. Proteins were precipitated by adding saturated NaCl, mixing on rocking platform for 5

minutes and centrifuging at 15000 g for 10 minutes at 4°C. Genomic DNA was precipitated adding 2-propanol and washed with 70 % ethanol. DNA was resuspended in TE and analysed by PCR. DNA that was analysed by Southern blot was further purified by two rounds of phenol-chloroform extraction and ethanol precipitation.

The genotype of mice was also analysed by PCR. Heterozygous animals were distinguished from wild type by the presence of the selectable gene using a pair of primers synthesised on the sequence of *LacZ* gene. Sequences of the oligonucleotides:

lacZdir 5'-TCACGAGCATCATCCTC-3'

lacZrev 5'-CGACAGATTTGATCCAGC-3'

PCR conditions: thirty-five cycles of denaturation (30 s at 94°C), of annealing (45 s at 56°C), and extension (30 s at 72°C).

To discriminate knock-out mice from heterozygous we looked for the absence of intron 3 that was amplified using the same conditions described in chapter 4.4.

4.16 Heteroduplex analysis

500 µl of blood were taken from tail of each mouse to analyse. To isolate T lymphocytes, the blood was diluted 6-8 fold in PBS and was layered on a Ficoll cushion and centrifuged at 800 g for 20 minutes in a swing-out rotor with low acceleration and without brake. The ring containing T lymphocytes was recovered and washed twice on PBS. Total RNA was extracted from T lymphocytes using the procedure described by Chomczynski and Sacchi (Chomczynski and Sacchi, 1987). 1 µg of RNA were converted into cDNA in a total volume of 20 µl with the following conditions: 50 mM Tris-HCl pH8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT. 0.5 mM dNTP, 30 units of RNase Inhibitor, 50 ng of oligo (dT) primer and 200 units of MMLV reverse transcriptase (BRL, Gaithersburg, MD. USA). RNA was preheated at 65°C for 5 minutes, chilled on ice for 1 minute, added to the mixture and incubated at 37°C for 1 hour. The reverse transcriptase was heat inactivated at 65°C for 10 minutes. The presence

of cDNAs was analysed by PCR amplification a 450 bp long fragment on the constant ß chain of the TCR. Sequences of the primers synthesised on the constant chain of TCR

upCßex1 5'-GAAATGTGACTCCACCCAAG-3'

dwCßex3 5'-AGGATCTCATAGAGGATGGT-3'

PCR conditions: 35 cycles of denaturation (30 s at 94°C), annealing (45 s at 55°C) and extension (30 s at 72°C).

1/20 of each cDNA sample was then amplified with 100 ng of each Vß specific oligonucleotide and 100 ng of dwCßex1 (5'-TCCTTGTAGGCCTGAGGGTCC-3') in a PCR mixture that contains 1X DNA polymerase buffer, 0.2 mM dNTPs, 0.6 units of DNA polymerase from Finnzyme. PCR conditions: 35 cycles of denaturation (30 s at 94°C), annealing (45 s at 55°C) and extension (30 s at 72°C).

Sequences of all the primers synthesised on the Vß families of TCR:

Vß2 5'-GAAGAATTCCCAGTATCCCTG-3'

Vß5.1 5'-GTATTCCCATCTCTGGACATCT-3'

Vß5.2 5'-GTATTCCCATCTCTGGACATAG-3'

Vß6 5'-CAATCATGATACAATGTACTGG-3'

Vß8.1 5'-AAAGGTGACATTGAGCTGTCAC-3'

Vß8.2 5'-AAAGGTGACATTGAGCTGTAAT-3'

Vß8.3 5'-GAAACGTGACATTGAGCTGTCG-3'

2/3 of the PCR products were denatured at 95°C for 5 min and then kept at 50°C for 1 hour. This temperature is permissive for the annealing between either homologous (homoduplex) or heterologous DNA strands, which have the same Vβ but differ in the J and the N regions (heteroduplex). The annealed PCR samples were separated at 4°C on a 12% acrylammide gel in 0.5X TBE for 16 hours at 10 mM. Gels were stained with ethidium bromide for 20 min in 0.5X TBE and photographed under UV light. Gels were blotted onto nylon membranes (GeneScreen *Plus*) using 20X SSC buffer for 48 hours. Transferred DNA was denatured and fixed to the membrane using 3 MM paper soaked

with 0.4M NaOH for 20 min. The filter was hybridised with a primer synthesised on exon 1 of the constant ß chain (dwCßex15'-TCCTTGTAGGCCTGAGGGTCC-3'), ^{32}PγATP labelled, in a solution containing 6X SSC, 1% BLOTTO, 5 mM EDTA, and 0.1% SDS. 1 at 37°C. The filter was washed twice in a solution containing 6X SSC, 0.1% SDS at 50°C for 30 min and then was exposed to X-ray film.

4.17 *In situ* hybridisation, immunohistochemistry and histological staining

Embryos from different development stages were dissected from pregnant mothers, for *in situ* hybridisation. Different adult tissues were dissected from knock-out and wild type mice, for histological analysis. Generally, embryos and organs were washed in phosphate buffered saline (PBS) and fixed O/N at 4 °C in freshly prepared 4 % paraformaldehyde in PBS. Tissues were then washed in PBS and dehydrated through one change each of 50 %, 70 %, 85 %, 95 % ethanol in 0.9 % NaCl and two changes in 100 % ethanol. Tissues were then embedded in paraffin, through two changes at room temperature in xylene for 1 h each, one change at 60 °C in 50 % xylene-paraffin O/N and two or three changes at 60 °C in 100% paraffin. Paraffin blocks were cut into 7 µm tissue sections, which were then mounted on subbed slides to be further processed for *in situ* hybridisation and histological analysis.

For *in situ* hybridisation, the specimens were dewaxed and rehydrated through the ethanol scale and hybridised O/N at 55 °C with a riboprobe obtained by *in vitro* transcription of the HMG2 cDNA cloned into the pBlueScript II KS(+) plasmid.

Transcription bf (5X, Boehringer)	4 µl
DTT (100 mM)	2 µl
2.5 mM ATP, CTP, GTP	4 µl
UTP (100 µM)	2.4 µl
RNase inhibitor (Promega)	20 units

DNA	1.5 µg
(^{35}S)UTPα (100mCi)	5 µl
T3 or T7 RNA pol (Boehringer)	15-20 units
H_2O	to 40 µl

The hybridisation mix contained 50 % formamide, 0.3 M NaCl, 20 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 10 % dextran sulphate, 1X Denhardt's solution, 0.5 mg/ml of yeast tRNA and the probe at a final concentration of 5-10 X10^4 cpm/ml. The slides were washed at 60 °C for 1 h with a solution containing 50 % formamide, 2X SSC, 10 mM DTT. The slides were digested with RNase, dehydrated and exposed against film. After exposure the slides were dipped in autoradiographic emulsion, developed and analysed.

Non radioactive *in situ* hybridisation was performed in adult testes sections. The same probe used for the radioactive *in situ* hybridisation was labelled with digoxigenin-UTP by *in vitro* transcription with the proper RNA polymerases (DIG RNA Labelling Kit from Boehringer Mannheim):

Transcription bf (10X, Boehringer)	2 µl
nucleotides mix (10mM ATP, CTP, GTP;	
6.5mM UTP; 3.5mM DIG-11-UTP	4 µl
RNase inhibitor (Promega)	20 units
DNA	1.5 µg
T3 or T7 RNA pol (Boehringer)	15-20 units
H_2O	to 20 µl

The sections were hybridised and washed using the same hybridisation conditions used for the radioactive *in situ* hybridisation. To reveal the DIG-labelled RNA probe, the sections were first incubated O/N with an anti-digoxigenin alkaline phosphatase conjugate antibody, and then were incubated with a colour-substrate solution containing 5-bromo-4-cloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium salt (NBT). The alkaline

phosphatase catalyses a reaction with BCIP and NBT that produce an insoluble blue/purple precipitate, which visualises hybrid molecules.

For histological analysis, the tissue sections were stained with hematoxylin-eosin. The slides were rehydrated through the ethanol scale and were incubated for 30 seconds with hematoxylin stain. The sections were then rinsed two times in tap water and stained for 50 seconds with eosin. To mount the slides, tissue sections were dipped eight times in the same 95 % ethanol solution, and eight times in 100 % ethanol. After a step of 5 minutes in xylene, the slides were stably mounted with DPX mounting medium.

For immunohystochemistry analysis, tissue sections were deparaffinised and hydrated through xylene and graded alcohol series. To quench endogenous peroxidase activity, the sections were incubated for 30 minutes in 0.3% H_2O_2 in methanol and then rinsed in PBS. To favour the recognition of antigens by the antibody, the slides were dipped three times for 5 minutes in boiling TE solution (10 mM Tris-HCl pH 8, 1 mM EDTA). The sections were then incubated for 1 hour at 4°C with normal blocking serum and then incubated O/N with 2.35 µg/ml of the polyclonal rabbit anti-HMG2 antibody from PharMingen. The sections were then incubated at room temperature for two hours with 2 µg/ml of the biotinylated anti-rabbit antibody. The sections were then incubated with a preformed Avidin and Biotinylated horse-radish peroxydase macromolecular Complex (VECTASTAIN Elite ABC kit from VECTOR), followed by an incubation with the peroxydase substrate until staining developed. As chromogen was used the 3-amino-9-ethyl carbazole (AEC), that is a peroxidase substrate that produces a red reaction product in the sections. The slides were then mounted in a aqueous mounting medium.

Tissues from heterozygous and knock-out mice were analysed for the expression of the ß-galactosidase (ß-gal). The construct for the gene targeting of *hmg2* carries the selectable marker fused to *LacZ* gene whose expression is under the control of the endogenous promoter of *hmg2*. ß-gal distribution was analysed by ß-gal staining assay. Tissues were dissected from +/- mice, washed in PBS and fixed in 4 % fresh

paraformaldeyde. Tissues were then incubated O/N at 37°C in the ß-gal staining solution containing 1mg/ml of 3-indolyl-ß-D-galactopyranoside (X-GAL) substrate in combination with 5mM potassium hexaferrocyanide and 5mM potassium hexaferricyanide in 1X PBS. Stained tissues were then paraffin embedded, cut into 7 µm tissue-sections, mounted onto glass slides, and analysed by light microscopy.

4.18 Apoptosis detection in tissue sections

Apoptosis was detected measuring the fragmented DNA of apoptotic cells by catalytically incorporating fluorescein-12-dUTP at 3'OH DNA ends using the enzyme terminal deoxynucleotidyl tranferase (TdT), which forms a polymeric tail (TUNEL assay, TdT-mediated dUTP Nick-End Labelling). Paraffin embedded testis sections were deparaffinised and rehydrated through graded ethanol washes and fixed again in 4% methanol-free formaldehyde solution in PBS. Sections were permeabilised for 8 minutes with 20 µg/ml of proteinase K and fixed again with 4% methanol-free formaldehyde. Slides were incubated with 1X equilibration buffer (200mM potassium cacodylate pH 6.6, 25 mM Tris-HCl pH 6.6, 0.2 mM DTT, 0.25 mg/ml BSA, and 2.5 mM cobalt chloride) and then incubated at 37°C for 1 hour with 50 µl of a solution containing 45 µl of equilibration buffer, 5 µl of nucleotide mix (50 µM fluorescein-12-dUTP, 100 µM dATP, 10 mM Tris-HCl pH 7.6, and 1 mM EDTA), and 25 units of TdT enzyme. The slides were washed three times with 2X SSC and once with PBS, the slides were then mounted with an aqueous mounting medium containing 4',6-diamidino-2-phenylindole (DAPI) and analysed at the fluorescence microscope.

4.19 *In vitro* fertilisation assay

Sperm was isolated from the epididymis of sacrificed *hmg2* knock-out male mice. The epididymis was isolated from the testis and from excess fat. The vas deferens was cut close to the cauda epididymis and sperm was recovered into 500 µl of pregassed M16

medium (SIGMA). The sperm was then incubated for 1 hour and half at 37°C for capacitation. Only the sperm from the cauda epididymis is suitable for *in vitro* fertilisation.

CD1 female mice were induced to superovulate by injecting an aqueous solution of pregnant mare serum gonadotropin (PMSG) and of human chorionic gonadotropin (hCG), respectively 58 and 12 hours before eggs recovery. Mice were sacrificed, the oviduct was dissected and the eggs were recovered in M16 medium. The eggs were then released from the cumulus masses by adding to the M16 medium hyaluronidase at a final concentration of 300 mg/ml. After this treatment the eggs were transferred with a sterile transfer pipette into 1000 µl of pregassed M16 medium and then were incubated for 4 hours with 100 µl of sperm, corresponding to 1 to $2X10^6$ sperm/ml. We consider fertilised the eggs that shown extrusion of the second polar body.

4.20 Isolation of Mouse Embryonic Fibroblasts

Mouse Embryonic Fibroblasts (MEFs) were isolated from the skin of +/+, +/- and -/- embryos at development stage E18. The embryos were dissected from a +/- pregnant mother mated with a +/- male. The skin and the tail of each embryo were recovered. The tail was used to extract DNA to genotype the embryos. The skin of each embryo was washed three times in PBS and digested at 37°C for thirty minutes in 1X trypsin-EDTA (GIBCO). The supernatant was recovered and centrifuged at 1500 rpm for 5 minutes. The pellet was resuspended in three ml of DMEM supplemented with 10% of FBS and antibiotics and the cells were plated into a 1.5 cm petri dish. A second digestion of 30-45 minutes was done with the rest of the skin, in order to recover more fibroblasts. After 5 hours the medium of the cells was changed. The cells were grown for several passages without immortalisation.

REFERENCES

Affolter, M., Percival-Smith, A., Muller, M., Billeter, M., Qian, Y. Q., Otting, G., Wuthrich, K. and Gehring, W. J. (1991). Similarities between the homeodomain and the Hin recombinase DNA- binding domain. Cell *64*, 879-80.

Agrawal, A., Eastman, Q. M. and Schatz, D. G. (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature *394*, 744-51.

Agrawal, A. and Schatz, D. G. (1997). RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. Cell *89*, 43-53.

Aidinis, V., Bonaldi, T., Beltrame, M., Santagata, S., Bianchi, M. E. and Spanopoulou, E. (1999). The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. Mol Cell Biol *19*, 6532-6542.

Anthony, C. T., Danzo, B. J. and Orgebin-Crist, M. C. (1984). Investigations on the relationship between sperm fertilizing ability and androgen-binding protein in the restricted rat. Endocrinology *114*, 1413-8.

Bellve, A. R., Cavicchia, J. C., Millette, C. F., O'Brien, D. A., Bhatnagar, Y. M. and Dym, M. (1977). Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. J Cell Biol *74*, 68-85.

Bergman, Y., Rice, D., Grosschedl, R. and Baltimore, D. (1984). Two regulatory elements for immunoglobulin kappa light chain gene expression. Proc Natl Acad Sci U S A *81*, 7041-5.

Bernues, J., Querol, E., Martinez, P., Barris, A., Espel, E. and Lloberas, J. (1983). Detection by chemical cross-linking of interaction between high mobility group protein 1 and histone oligomers in free solution. J Biol Chem *258*, 11020-4.

Bianchi, M. E. (1988). Interaction of a protein from rat liver nuclei with cruciform DNA. EMBO J. *7*, 843-849.

Bianchi, M. E., Beltrame, M. and Paonessa, G. (1989). Specific recognition of cruciform DNA by nuclear protein HMG1. Science *243*, 1056-1059.

Bonne-Andrea, C., Harper, F., Sobczak, J. and De Recondo, A. M. (1984). The role of HMG1 protein in nucleosome assembly and in chromatin replication. Adv Exp Med Biol *179*, 479-88.

Boonyaratanakornkit, V., Melvin, V., Prendergast, P., Altmann, M., Ronfani, L., Bianchi, M. E., Taraseviciene, L., Nordeen, S. K., Allegretto, E. A. and Edwards, D. P. (1998). High-mobility group chromatin proteins 1 and 2 functionally interact with steroid hormone receptors to enhance their DNA binding in vitro and transcriptional activity in mammalian cells. Mol Cell Biol *18*, 4471-87.

Brook, C. G., Wagner, H., Zachmann, M., Prader, A., Armendares, S., Frenk, S., Aleman, P., Najjar, S. S., Slim, M. S., Genton, N. and Bozic, C. (1973). Familial

occurrence of persistent Mullerian structures in otherwise normal males. Br Med J *1*, 771-3.

Bucci, L. R., Brock, W. A., Goldknopf, I. L. and Meistrich, M. L. (1984). Characterization of high mobility group protein levels during spermatogenesis in the rat. J Biol Chem *259*, 8840-6.

Bucci, L. R., Brock, W. A. and Meistrich, M. L. (1985). Heterogeneity of high-mobility-group protein 2. Enrichment of a rapidly migrating form in testis. Biochem J *229*, 233-40.

Bustin, M., Lehn, D. A. and Landsman, D. (1990). Structural features of the HMG chromosomal proteins and their genes. Biochim Biophys Acta *1049*, 231-243.

Calogero, S., Grassi, F., Aguzzi, A., Voigtlander, T., Ferrier, P., Ferrari, S. and Bianchi, M. E. (1999). The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. Nat Genet *22*, 276-80.

Canaple, L., Decoville, M., Leng, M. and Locker, D. (1997). The Drosophila DSP1 gene encoding an HMG 1-like protein: genomic organization, evolutionary conservation and expression. Gene *184*, 285-90.

Carballo, M., Puigdomenech, P. and Palau, J. (1983). DNA and histone H1 interact with different domains of HMG 1 and 2 proteins. EMBO J *2*, 1759-64.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. Anal Biochem *162*, 156-9.

Clarke, H. J., Oblin, C. and Bustin, M. (1992). Developmental regulation of chromatin composition during mouse embryogenesis: somatic histone H1 is first detectable at the 4-cell stage. Development *115*, 791-9.

Clincke, G. and Wauquier, A. (1979). Metrazol-produced impairment of passive avoidance retention specifically antagonized by anti-petit mal drugs. Psychopharmacology *66*, 243-6.

Crippa, M. P., Alfonso, P. J. and Bustin, M. (1992). Nucleosome core binding region of chromosomal protein HMG-17 acts as an independent functional domain. J Mol Biol *228*, 442-9.

Davis, D. L. and Burch, J. B. E. (1992). Isolation of a chicken *HMG2* cDNA clone and evidence for an HMG2-specific 3'-untranslated region. Gene

Difilippantonio, M. J., McMahan, C. J., Eastman, Q. M., Spanopoulou, E. and Schatz, D. G. (1996). RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. Cell *87*, 253-62.

Dimitrov, S., Almouzni, G., Dasso, M. and Wolffe, A. P. (1993). Chromatin transitions during early Xenopus embryogenesis: changes in histone H4 acetylation and in linker histone type. Dev Biol *160*, 214-27.

Dimitrov, S., Dasso, M. C. and Wolffe, A. P. (1994). Remodeling sperm chromatin in Xenopus laevis egg extracts: the role of core histone phosphorylation and linker histone B4 in chromatin assembly. J Cell Biol *126*, 591-601.

Drlica, K. and Rouviere-Yaniv, J. (1987). Histonelike proteins of bacteria. Microbiol Rev *51*, 301-19.

Eddy, E. M., Washburn, T. F., Bunch, D. O., Goulding, E. H., Gladen, B. C., Lubahn, D. B. and Korach, K. S. (1996). Targeted disruption of the estrogen receptor gene in male mice causes alteration of spermatogenesis and infertility. Endocrinology *137*, 4796-805.

Einck, L. and Bustin, M. (1985). The intracellular distribution and function of the High Mobility Group chromosomal proteins. Exp Cell Res *156*, 295-310.

Falciola, L., Spada, F., Calogero, S., Langst, G., Voit, R., Grummt, I. and Bianchi, M. E. (1997). High mobility group 1 protein is not stably associated with the chromosomes of somatic cells. J Cell Biol *137*, 19-26.

Ferrari, S., Ronfani, L., Calogero, S. and Bianchi, M. E. (1994). The mouse gene coding for high mobility group 1 protein (HMG1). J Biol Chem *269*, 28803-8.

Giese, K., Cox, J. and Grosschedl, R. (1992). The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. Cell *69*, 185-95.

Goodwin, G. H., Woodhead, L. and Johns, E. W. (1977). The presence of high mobility group non-histone chromatin proteins in isolated nucleosomes. FEBS Lett *73*, 85-8.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Muensterberg, A., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature *346*, 245-250.

Hammond, G. L. (1990). Molecular properties of corticosteroid binding globulin and the sex- steroid binding proteins. Endocr Rev *11*, 65-79.

Hayashi, S. and Scott, M. P. (1990). What determines the specificity of action of Drosophila homeodomain proteins?. Cell *63*, 883-94.

He, W. W., Fischer, L. M., Sun, S., Bilhartz, D. L., Zhu, X. P., Young, C. Y., Kelley, D. B. and Tindall, D. J. (1990). Molecular cloning of androgen receptors from divergent species with a polymerase chain reaction technique: complete cDNA sequence of the mouse androgen receptor and isolation of androgen receptor cDNA probes from dog, guinea pig and clawed frog. Biochem Biophys Res Commun *171*, 697-704.

Hutson, J. M. (1986). Testicular feminization: a model for testicular descent in mice and men. J Pediatr Surg *21*, 195-8.

Hutson, J. M. and Donahoe, P. K. (1986). The hormonal control of testicular descent. Endocr Rev *7*, 270-83.

Jantzen, H.-M., Admon, A., Bell, S. P. and Tjian, R. (1990). Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature *344*, 830-836.

Jenuwein, T. and Grosschedl, R. (1991). Complex pattern of immunoglobulin mu gene expression in normal and transgenic mice: nonoverlapping regulatory sequences govern distinct tissue specificities. Genes Dev *5*, 932-43.

Krumlauf, R. (1994). Hox genes in vertebrate development. Cell *78*, 191-201.

Lavoie, B. D. and Chaconas, G. (1993). Site-specific HU binding in the Mu transpososome: conversion of a sequence-independent DNA-binding protein into a chemical nuclease. Genes Dev *7*, 2510-9.

Lavoie, B. D. and Chaconas, G. (1994). A second high affinity HU binding site in the phage Mu transpososome. J Biol Chem *269*, 15571-6.

Lewis, S. M. (1994). The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. Adv Immunol *56*, 27-150.

Lilley, D. M. J. (1990a). Genetic recombination: four-way DNA junctions and their resolution. Curr Opin Cell Biol *2*, 464-7.

Lilley, D. M. J. (1990b). The structure of the four-way helical junction in DNA, and its role in genetic recombination. Berlin. Springer Verlag

Manak, J. R. and Scott, M. P. (1994). A class act: conservation of homeodomain protein functions. Dev Suppl 61-77.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1989). Molecular cloning: a laboratory manual (Cold Spring Harbour Laboratory Press, Cold Spring Harbour).

Nef, S. and Parada, L. F. (1999). Cryptorchidism in mice mutant for *Insl3*. Nat Genet *22*, 295-299.

Ner, S. S. and Travers, A. A. (1994). HMG-D, the Drosophila melanogaster homologue of HMG 1 protein, is associated with early embryonic chromatin in the absence of histone H1. EMBO J *13*, 1817-22.

Nightingale, K., Dimitrov, S., Reeves, R. and Wolffe, A. P. (1996). Evidence for a shared structural role for HMG1 and linker histones B4 and H1 in organizing chromatin. EMBO J *15*, 548-61.

Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O. and Lippard, S. J. (1999). Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. Nature *399*, 708-12.

Oñate, S. A., Prendergast, P., Wagner, J. P., Nissen, M., Reeves, R., Pettijohn, D. E. and Edwards, D. P. (1994). The DNA bending protein HMG-1 enhances progesterone receptor binding to its target DNA sequences. Mol Cell Biol *14*, 3376-3391.

Perez-Palacios, G. and Jaffe, R. B. (1972). The syndrome of testicular feminization. Pediatr Clin North Am *19*, 653-67.

Pil, P. M., Chow, C. S. and Lippard, S. J. (1993). High-mobility-group 1 protein mediates DNA bending as determined by ring closures. Proc Natl Acad Sci U S A *90*, 9465-9.

Pil, P. M. and Lippard, S. J. (1992). Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. Science *256*, 234-236.

Pontiggia, A., Negri, A., Beltrame, M. and Bianchi, M. E. (1993). Protein HU binds specifically to kinked DNA. Mol Microbiol *7*, 343-350.

Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C. and Norman, D. G. (1993). Solution structure of a DNA-binding domain from HMG1. Nucleic Acids Res *21*, 3427-36.

Read, C. M., Cary, P. D., Preston, N. S., Lnenicek-Allen, M. and Crane-Robinson, C. (1994). The DNA sequence specificity of HMG boxes lies in the minor wing of the structure. EMBO J *13*, 5639-46.

Sawchuk, D. J., Weis-Garcia, F., Malik, S., Besmer, E., Bustin, M., Nussenzweig, M. C. and Cortes, P. (1997). V(D)J recombination: modulation of RAG1 and RAG2 cleavage activity on 12/23 substrates by whole cell extract and DNA-bending proteins. J Exp Med *185*, 2025-32.

Schatz, D. G., Oettinger, M. A. and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. Cell *59*, 1035-48.

Scovell, W. M., Muirhead, N. and Kroos, L. R. (1987). cis-Diamminedichloroplatinum(II) selectively cross-links high mobility group proteins 1 and 2 to DNA in micrococcal nuclease accessible regions of chromatin. Biochem Biophys Res Commun *142*, 826-35.

Seyedin, S. M. and Kistler, W. S. (1979). Levels of chromosomal protein high mobility group 2 parallel the proliferative activity of testis, skeletal muscle, and other organs. J Biol Chem *254*, 11264-71.

Sharman, A. C., Hay-Schmidt, A. and Holland, P. W. (1997). Cloning and analysis of an HMG gene from the lamprey Lampetra fluviatilis: gene duplication in vertebrate evolution. Gene *184*, 99-105.

Shirakawa, H. and Yoshida, M. (1992). Structure of a gene coding for human HMG2 protein. J Biol Chem *267*, 6641-6645.

Shirakawa, H. and Yoshida, M. (1995). Existence of a transcription factor for the human HMG2 gene positively related to the level of HMG2 mRNA in the cells. Biochemistry *34*, 2521-7.

Sinclair, A., Berta, P., Palmer, M. S., Hawkins, R. J., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A.-M., Lovell-Badge, R. and Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature *346*, 240-244.

Singh, J. and Dixon, G., H. (1990). High mobility group proteins 1 and 2 function as general class II transcription factors. Biochemistry *29*, 6295-6302.

Spada, F., Brunet, A., Mercier, Y., Renard, J. P., Bianchi, M. E. and Thompson, E. M. (1998). High mobility group 1 (HMG1) protein in mouse preimplantation embryos. Mech Dev *76*, 57-66.

Spanopoulou, E., Zaitseva, F., Wang, F. H., Santagata, S., Baltimore, D. and Panayotou, G. (1996). The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. Cell *87*, 263-76.

Staben, C. and Yanofski, C. (1990). Neurospora crassa a mating-type region. Proc. Natl. Acad. Sci. USA *87*, 4917-4921.

Stewart, M. (1993). Disinhibition of hippocampal pyramidal cells during the transition into theta rhythm. Exp Brain Res *93*, 1-5.

Stewart, M. and Wong, R. K. (1993). Intrinsic properties and evoked responses of guinea pig subicular neurons in vitro. J Neurophysiol *70*, 232-45.

Sugimoto, A., Iino, Y., Maeda, T., Watanabe, Y. and Yamamoto, M. (1991). Schizosaccharomyces pombe ste11+ encodes a transcription factor with an HMG motif that is a critical regulator of sexual development. Genes Dev *5*, 1990-9.

Thanos, D. and Maniatis, D. (1992). The high mobility group protein HMG-I(Y) is required for NF-kB-dependent virus induction of the human IFN-β gene. Cell *71*, 777-789.

Travers, A. A., Ner, S. S. and Churchill, M. E. (1994). DNA chaperones: a solution to a persistence problem? Cell *77*, 167-9.

Travis, A., Amsterdam, A., Belanger, C. and Grosschedl, R. (1991). LEF-1, a gene encoding a lymphoid-specific potein, with an HMG domain, regulates T-cell receptor a enhancer function. Genes Dev *5*, 880-894.

Tremethick, D. J. and Molloy, P. L. (1986). High mobility group proteins 1 and 2 stimulate transcription in vitro by RNA polymerase II and III. J. Biol. Chem. *261*, 6986-6992.

Tremethick, D. J. and Molloy, P. L. (1988). Effects of high mobility group proteins 1 and 2 on initiation and elongation of specific transcription by RNA polymerase II in vitro. Nucleic Acids Res *4*, 1471-1486.

Ura, K. and Wolffe, A. P. (1996). Reconstruction of transcriptionally active and silent chromatin. Methods Enzymol *274*, 257-71.

Vaccari, T., Beltrame, M., Ferrari, S. and Bianchi, M. E. (1998). Hmg4, a new member of the Hmg1/2 gene family. Genomics *49*, 247-52.

van de Wetering, M., Oosterwegel, M., Dooijes, D. and Clevers, H. (1991). Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. EMBO J. *10*, 123-132.

van Gent, D. C., Hiom, K., Paull, T. T. and Gellert, M. (1997). Stimulation of V(D)J cleavage by high mobility group proteins. EMBO J *16*, 2665-70.

Wack, A., Montagna, D., Dellabona, P. and Casorati, G. (1996). An improved PCR-heteroduplex method permits high-sensitivity detection of clonal expansions in complex T cell populations. J Immunol Methods *196*, 181-92.

Waga, S., Mizuno, S. and Yoshida, M. (1990). Chromosomal protein HMG1 removes the transcriptional block caused by the cruciform in supercoiled DNA. J Biol Chem *265*, 19424-8.

Wagner, B. L., Norris, J. D., Knotts, T. A., Weigel, N. L. and McDonnell, D. P. (1998). The nuclear corepressors NCoR and SMRT are key regulators of both ligand- and 8-bromo-cyclic AMP-dependent transcriptional activity of the human progesterone receptor. Mol Cell Biol *18*, 1369-78.

Wanschura, S., Schoenmakers, E. F., Huysmans, C., Bartnitzke, S., Van de Ven, W. J. and Bullrdiek, J. (1996). Mapping of the human HMG2 gene to 4q31. Genomics *31*, 264-5.

Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R., Laue, E. D. and Thomas, J. O. (1993). Structure of the HMG box motif in the B-domain of HMG1. EMBO J *12*, 1311-9.

Werner, M. H., Bianchi, M. E., Gronenborn, A. M. and Clore, G. M. (1995). NMR spectroscopic analysis of the DNA conformation induced by the human testis determining factor SRY. Biochemistry *34*, 11998-2004.

Wirth, T., Staudt, L. and Baltimore, D. (1987). An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature *329*, 174-8.

Yamazaki, F., Nagatsuka, Y., Shirakawa, H. and Yoshida, M. (1995). Repression of cell cycle progression by antisense HMG2 RNA. Biochem Biophys Res Commun *210*, 1045-51.

Zappavigna, V., Falciola, L., Citterich, M. H., Mavilio, F. and Bianchi, M. E. (1996). HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. EMBO J *15*, 4981-91.

Zappavigna, V., Renucci, A., Izpisua-Belmonte, J. C., Urier, G., Peschle, C. and Duboule, D. (1991). HOX4 genes encode transcription factors with potential auto- and cross- regulatory capacities. EMBO J *10*, 4177-87.

Zwilling, S., Konig, H. and Wirth, T. (1995). High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. EMBO J *14*, 1198-208.