Functional characterisation of a HOX/PBX transcriptional complex

Thesis

Giuliana Di Rocco

FUNCTIONAL CHARACTERIZATION OF A HOX/PBX TRANSCRIPTIONAL COMPLEX

Thesis submitted in partial fulfilment of the requirements of the Open University for the degree of Doctor of Philosophy in Molecular and Cellular Biology

July 1998

DIBIT
Department of Biological and Technological Research - Milan, Italy

DECLARATION

This Thesis has been composed by myself and has not been used in any previous application for a degree. The results presented here were obtained by myself. All sources of information are acknowledged by means of reference.
Some of the work presented in this thesis has been published in:
Di Rocco G., Mavilio F., Zappavigna V. (1997). "Functional dissection of a transcriptionally active, target-specific Hox/Pbx complex". EMBO J. *16*, 3644-3654.

ACKNOWLEDGMENTS

First I would like to thank Fulvio Mavilio very much for letting me join his group and giving me the opportunity to carry out a Ph.D. in his lab. I am very grateful for his support and guidance throughout the project.

I would also especially like to thank Vincenzo Zappavigna for all the discussions and criticisms during the time we worked together. His help and advices made these years a very satisfying and fruitful experience.

Many thanks as well to all the members of the Gene Expression Laboratory at DIBIT in Milan who encouraged me and made the lab a really pleasant and enjoyable place.

Finally thank you very much to Prof. Robb Krumlauf for his interest and supervision during this Thesis.

Last, but certainly not least, I am greatly endebted to my parents for their continuous help and support.

ABSTRACT

Hox genes control cell fates and specify regional identities during vertebrate development. Hox proteins relaxed DNA-binding selectivity *in vitro*, suggests that Hox specificity of action is achieved *in vivo* through the action of transcriptional cofactors. Members of the EXD/PBX family of homeoproteins have been proposed for such a role on the basis of genetic and biochemical evidence. In this work we show that the human Pbx1 and HOXB1 proteins can cooperatively activate transcription through a genetically charaterized Hox responsive element, the autoregulatory enhancer of the mouse *Hoxb-1* gene (b-1ARE), which directs the spatially restricted expression of *Hoxb-1* in the fourth rhombomere during hindbrain development. On the b-1ARE, only a restricted subset of HOX proteins is able to bind cooperatively and activate transcription. Selective recognition of the b-1ARE is mediated by the N-terminal region of the HOX homeodomain. The DNA binding and protein-protein interaction functions of HOXB1 and Pbx1 are all necessary for the assembly of a transcriptionally active complex on the b-1ARE. Functional dissection of the complex allowed the localization of the main activation domain in the HOXB1 N-terminal region, and of an additional one in the C-terminal region of Pbx1, which is absent in one of its two alternative splicing isoforms. The transcriptional activity of HOX/Pbx1 complexes on the b-1ARE element is maximal and further restricted in embryonal carcinoma (EC) cells compared to other cell lines. Mutational analysis shows that an octamer-binding protein consensus site on the b-1ARE contributes to the activity of the enhancer in transfected cells and it is bound *in vitro* by octamer-like proteins from both EC cells and embryonic extracts.

CONTENTS

ILLUSTRATIONS

Chapter 1

Chapter 2

Chapter 1

INTRODUCTION

Homeobox genes encode DNA binding proteins that regulate gene expression and control various aspects of morphogenesis and cell differentiation. These genes contain a common sequence element of 180 bp, the homeobox, which was first discovered in the fruitfly *Drosophila melanogaster.* Subsequently, the homeobox was shown to occur in all metazoa ranging from sponges to vertebrates and also in plants and fungi, and has thus been evolutionary conserved throughout the three kingdoms of multicellular organisms (Gehring, 1994). The homeobox encodes a 60-aminoacid homeodomain that is responsible for the DNA binding of homeodomain containing proteins. In vertebrates, about 170 different homeobox genes have already been cloned and it has been estimated that more that 0.2% of the genes present in a genome may possess an homeobox (Stein *et al.* 1996).

Homeobox-containing genes can be conveniently divided into two subfamilies: 1) the clustered class of Hox genes also known as "homeotic" genes, which are implied in the specification of regional identity along the antero-posterior axis of the embryo and 2) the non clustered or divergent homeobox genes; the latter are scattered throughout the genome and fall into a number of groups based on sequence similarities as for example the vertebrate *pax, Msx, Emx, Otx* genes named after their homologs in the fly, i.e. the *paired, muscle segment homeobox, empty spiracles* and *orthodenticle* genes, respectively. These groups are remarkably different in their functions which are, in at least some

cases, conserved across vast evolutionary distances (Gehring, 1994; Manak *et al.* 1996).

1.1 THE CLUSTERED CLASS OF HOX GENES.

In *Drosophila* there are eight Hox genes grouped in two separate clusters, the Antennapedia (Ant-C) and Bithorax (Bx-C) Complexes, on the same chromosome. These genes include *labial* (*lab*), *proboscipedia* (*pb*), *deformed* (*Dfd*), *sex combs reduced* (*scr*) and *Antennapedia* (*Antp*) in the *Antennapedia* complex and *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *abdominal-B* (*abd-B*) in the *Bithorax* complex. They are collectively referred to as the *Drosophila* Homeotic Complex (HOM-C) (McGinnis and Krumaluf, 1992).

The normal function of Hox genes is to assign distinct positional identities to cells in different regions along the antero-posterior axis. This is thought to happen through the activation of batteries of target "realisator" genes that specify the properties of a particular tissue or organ primordia (Lawrence and Morata, 1994). The name "homeotic" derives from the Greek word "homeo" meaning "alike" and was originally used for *Drosophila* mutants in which one of the insect's body segment was transformed into the likeness of another. For example, loss of function mutants of the Ubx gene lead to the transformation of the third thoracic segment (T3) carrying halteres (small balancers) toward a second thoracic segment (T2) with wings, thus generating four-winged flies. On the other hand, gain of function mutants of Ubx lead to the transformation of the wings into a second pair of halteres. The conclusion from this is that Ubx is required for promoting haltere and suppressing wing development (Lewis, 1978). Similarly, ectopic expression of the Antp gene resulting from a spontaneous chromosomal inversion placing the protein coding region of this

gene under the control of a heterologous promoter, leads to the conversion of the antenna into mesothoracic legs (Struhl, 1981).

The mammalian Hox genes are defined by virtue of their homology with the genes of the homeotic complex (HOM-C) of *Drosophila*. Analysis of mouse and human Hox genes indicates that there are at least 39 of them organized in four clusters, named from A to D, each localized on a different chromosome. On the basis of sequence similarities and relative position in the complex, the individual Hox genes within the different clusters can be aligned with each other and with genes of the *Drosophila* HOM-C cluster. These similarities suggest that the four mammalian clusters probably arose from a single ancestral complex by expansion of the cluster through lateral gene duplication and by duplication of the clusters by chromosomal duplication or polyploidization. The mammalian Hox genes on a single cluster are numbered from 1 to 13 starting from the 3' end of the cluster. Genes with an equivalent position on each cluster form what is called a "paralogous group" (Fig. 1.1).

Both the timing and the spatial domains of expression of Hox gene are related to their position on the chromosome: genes that are located at the 3' end of the cluster are expressed earlier and most anteriorly, while genes located at the 5' end of the cluster are expressed later and in more posterior regions. This spatial and temporal correspondence of expression patterns with the chromosomal localization, termed "colinearity", was described originally in *Drosophila*, and has since been observed in all animals exhibiting an antero-posterior axial polarity (Douboule and Morata, 1994; Krumlauf, 1994).

Therefore, the evolutionary conservation not only concerns the homeodomain, but also the homeobox gene organization: both the DNA sequence and the expression patterns are conserved between paralogous genes of the different clusters, which underlines the importance of these genes in development and evolution.

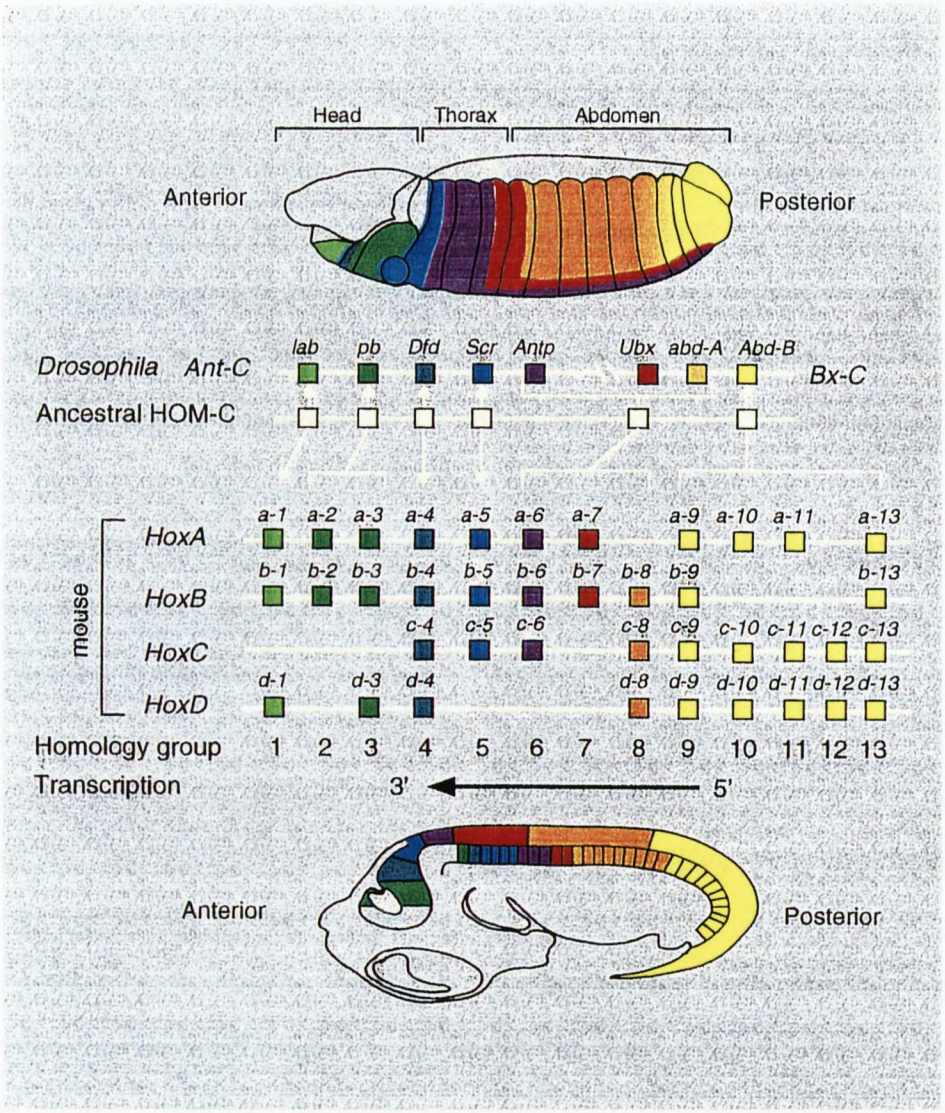


Figure 1.1. Genomic organization and colinear expression patterns of *Drosophila* HOM genes and murine Hox genes. The schemes of the *Drosophila Antennapedia* (*Ant-C*) and *Bithorax* (*Bx-C*) complexes, the four murine *Hox* complexes and a hypothetical ancestral HOM complex are shown with their possible phylogenetic relationships. Each gene is represented by a coloured box. The successive antero-posterior expression domains of HOM/*Hox* genes are schematized in a *Drosophila* embryo (upper drawing) and in the central nervous system and prevertebrae of a mid-gestation mouse embryo (lower drawing). The partial overlap between HOM gene transcripts in thoracic and abdominal segments of the fly embryo are indicated. The overlapping expression domains of murine *Hox* genes towards posterior region of the embryo could not be represented; hence, each colour is meant to show the anterior-most expression domain of a given subfamily. Consistent with their expression domains, *Hox* genes belonging to paralogous groups 1-4 primarily control the development of the branchial area and of the rhomboencephalon, central *Hox* genes of groups 5-8 control the thoracic portion of the body, whereas those one belonging to groups 9-13 control the lumbo-sacral region. (From Favier and Dollé, 1997).

1.2 EXPRESSION AND FUNCTION OF MAMMALIAN HOX GENES.

Mammalian Hox genes are expressed in nearly every cell type. They are transcribed in limited region along the antero-posterior axis, like the fly genes, although in more substantially overlapping patterns. In mammalian embryos, the earliest expression of Hox genes can be detected at gastrulation. They are expressed in all three germ layers with overlapping domains extending from the caudal end of the embryo to a sharp anterior limit that is specific for each Hox gene. Fewer Hox genes are expressed in the anterior than in the posterior. The position of anterior boundaries along the axis coincides with the position of each gene along the cluster, according to the colinearity rule. This is particularly evident in the mouse hindbrain neuroectoderm, where anterior expression boundaries colocalize with transient segmental units known as rhombomeres, and in the somites, where anterior boundaries of genes from all paralogous groups mark the anterior margin of prevertebrae (Keynes and Krumlauf, 1994; Kessel and Gruss, 1991).

In addition to the evolutionary conservation in structure, organization and expression with respect to *Drosophila* homeotic genes, studies of mice with Hox gene mutations that generate either gain or loss of function, reveal that vertebrate Hox genes, like their *Drosophila* counterparts, direct regional embryonic development and are involved in anterior-posterior axial pattern formation. For instance, ectopic expression of *Hoxa-7* induces conversion from the normal seven cervical vertebrae to eight cervical vertebrae and is accompanied by variations in the most anterior vertebrae that suggest a posterior to anterior transformation (Kessel *et al.*, 1990). Similarly, when the *Hoxd-4* gene is put under the control of the *Hoxa-1* promoter, which drives its expression into an ectopic anterior domain, it generates a transformation of the

exoccipital bones into ectopic ossified structures resembling vertebrae neural arches (Lufkin *et al.*, 1992).

To date, numerous loss of function mutations by targeted gene disruption of mammalian Hox genes have been reported (reviewed in Favier and Dolle', 1997), and more than half Hox genes have been functionally inactivated. Many of these "knock-out" experiments show "homeotic" transformation towards a more anterior phenotype in the regions were the genes are normally expressed as, for example, conversion of the first lumbar vertebra to a thoracic one, producing a supernumerary 14th pair of ribs by *Hoxc-8* inactivation (Le Mouellic *et al.*, 1992), or transformation of the second to first pharyngeal arch identity by *Hoxa-2* inactivation (Rijli *et al.*, 1993; Gendron-Maguire *et al.*, 1993). Others cause defects in patterning of hindbrain and pharyngeal arches, such as defects in the generation of rhombomeres 4-5, disorganized cranial nerves and inner ear abnormalities in *Hoxa-1* null mutants (Mark *et al.*, 1993; Carpenter *et al.*, 1993; Dollé *et al.*, 1993), and selective facial nerve motor neurons deficiencies in *Hoxb-1* and *Hoxb-2* knock-out mice (Goddard *et al.*, 1996; Studer *et al.*, 1996; Barrow and Capecchi, 1996).

The results of gain and loss of function experiments show the existence of a sort of hierarchy for homeotic gene function in both mice and flies. In general, posterior Hox proteins are dominant over more anterior ones, a phenomenon which has been termed "posterior prevalence" (Douboule, 1994). Although there are exceptions to the rule, the morphogenetic program seems to be imparted by the most posterior Hox expressed at a certain axial level. All together these data provide evidence that a common Hox network has been maintained to impart morphologic identities to segmented structures in animal groups employing radically different developmental strategies, and that some aspects of gene function has been conserved since the divergence of arthropods and chordates. This phylogenetic conservation of function is strengthened by

the ability to rescue loss of function mutations in *Drosophila* by vertebrate Hox gene, such as the rescue of *Drosophila* null mutant *labial* by its chicken ortholog *Hoxb-1* (Lutz *et al.*, 1996).

Many Hox genes, and in particular the Abd-B-related paralogs 9-13, are expressed in the developing limb in overlapping domains, similarly to what happens in axial structures and consistent with a Hox role in the specification of the digit pattern. Indeed, knock-out experiments of Abd-B related genes resulted in size reduction, changes in shape and/or delayed ossification of limb skeletal elements (Favier *et al.*, 1996; Fromental-Ramain *et al.*, 1996; Davis and Capecchi, 1996; Zàkàny and Douboule, 1996). Several data show the involvement of these genes also in the development of the genito-urinary tract and the terminal part of the digestive tract (Hsieh-Li *et al.*, 1995; Benson *et al.*, 1996; Kondo *et al.*, 1996). This implies that the Hox system is required at different structural and morphogenetic levels during development, wherever there is the necessity to define "axial values", as in the vertebral column or in the limb.

The area of influence of clustered Hox genes in embryo development has an anterior limit coincident with spinal cord and hindbrain, the so called "posterior head". The "rostral head", comprising forebrain and head mesoderm, is patterned by additional homeobox genes not included in Hox complexes such as *Otx* and *Emx* families, homologous to *Drosophila otd* and *ems*, as well as an increasing number of other regulatory genes. At least some of these head homeobox genes may have once been located in a primordial homeotic complex. The head genes may have functions analogous to those of the Hox genes acting in the trunk, but the complexities of anterior structures make the regulatory role of these genes less clear (Bally-Cuif and Boncinelli, 1996).

1.3 SPECIFICITY OF ACTION OF HOMEODOMAIN PROTEINS.

As mentioned above, a primary function of homeodomain protein is the transcriptional control of specific target genes. Although homeobox genes share a very conserved DNA binding domain, each homeoprotein has its own characteristic effects on the development of embryo and adult, only in part due to its unique spatial and temporal expression pattern. Furthermore, different Hox proteins have different and specific effects on morphology when ectopically expressed in the same compartments. A simple model to explain how Hox proteins control cell fate is that they each regulate the transcription of a different (although possibly overlapping) set of downstream target genes. However, *in vitro*, most Hox proteins bind the same or very similar binding sites with similar affinities. Key questions are, therefore, to what degree different homeodomain proteins control different target genes, and how is the specificity of action of the proteins attained.

Most of what is known about homeobox gene binding specificity derives from *in vitro* studies with purified proteins. The crystal structure of three homeodomains (engrailed, MATα2 and even-skipped) and the solution structure of one homeodomain (Antp) complexed with DNA have been solved with atomic resolution (Kissinger *et al., 1990;* Wolberger *et al., 1991;* Billeter *et al., 1993*). The homeodomain consists of three well defined α-helices, a more flexible fourth helix and an unstructured N-terminal arm. Helices 2 and 3 form an helix-turn-helix motif virtually identical to those observed in various prokaryotic transcriptional repressors, except that helix 3 is elongated considerably by helix 4. DNA contacts are mediated by helix 3, the so called "recognition" helix, which sits in the major groove, and by the flexible amino-

terminal arm which establishes contacts with bases in the minor groove of the DNA (Gehring *et al.*, 1994) (Fig. 1.2).

A single homeodomain can recognize a variety of sequences *in vitro* with similar affinities, while different homeodomains can bind to the same sequence. As a consequence, the specificity of recognition is lower compared to other sequence specific DNA binding proteins. In contrast to bacterial helix-turn-helix proteins, the homeoproteins appear to bind as monomers, which may account for the observed lower binding affinities. One limitation in interpreting binding data from *in vitro* experiments is that the sequences used for binding are in most cases artificial, obtained from searches of large pseudorandom sequences, such as phage lambda genome. Many of them contain a TAAT motif, but the number and the variety of homeodomain bound sequences has increased so much that is very difficult to derive a distinct consensus sequence. It can however be summarized as a core of six base pairs: 5'- TNAT(G/T)(G/A) -3', where N can be any nucleotide.

Data concerning homeoproteins specificity *in vivo* have been obtained by ectopically expressing chimeric proteins both in mice and in *Drosophila*. The ubiquitous expression of individual Hox proteins, in fact, generates specific phenotypes that mimic the wild type functions of these proteins and which can be used as a measure of the *in vivo* function. In general all this studies bring to the conclusion that the specificity of Hox proteins resides mostly in the homeodomain plus some residue close to the homeodomain (reviewed in Mann, 1995).

Differences in the primary sequence of homeodomains, especially differences in the "contact" regions (i.e. recognition helix and N-terminal arm), results in subtle differences in their *in vitro* DNA binding preferences, and are also critical for their specificity *in vivo*. A single change at position 9 of helix 3 which makes critical contacts with bases in the major groove, for example, is

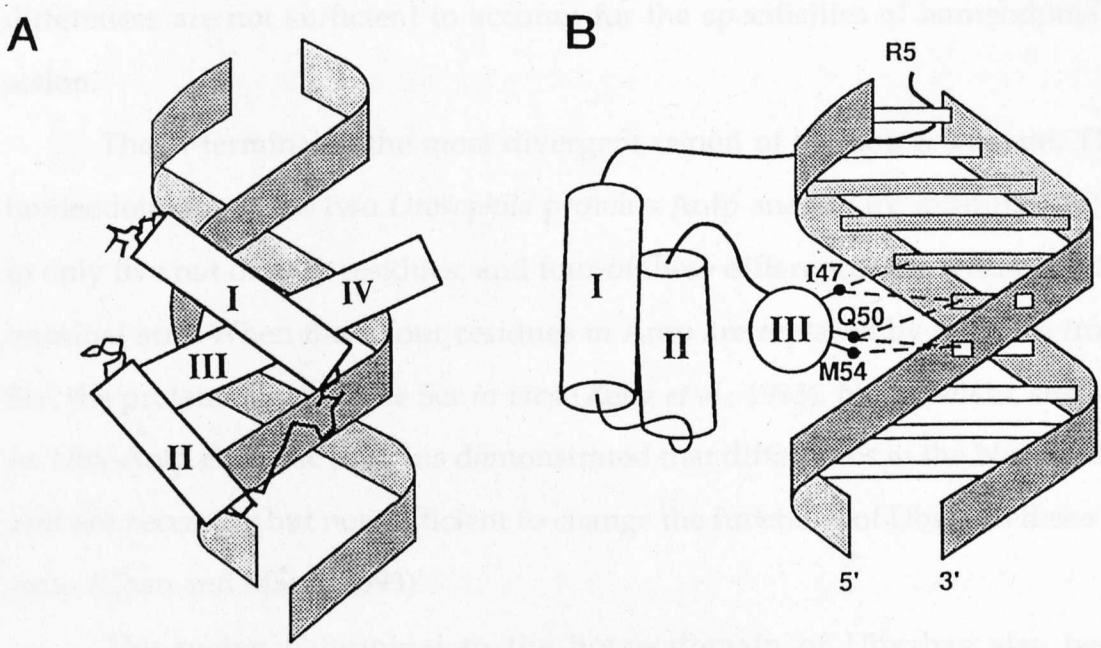


Figure 1.2. Schematic representation of the Homeodomain-DNA complex of Antp. (A) View perpendicular to the axis of the recognition helix (III, IV) located in the major groove. The a helices I, II and III, IV are indicated by cylinders. For the N-terminal arm (upper left), the loop between helices I and II and the turn of the helix-turn-helix motif (II-III) only the polypeptide backbone is drawn as a solid line. (B) View along the axis of the recognition helix (III). Aminoacid residues that establish contacts to specific bases are indicated. For residues in the recognition helix, these contacts, which are indicated by dotted lines, are in the major groove; for those in the N-terminal arm, they are in the minor groove of the DNA. The TAAT motif (see text) is shaded. (From Gehring *et al.*, 1994).

sufficient to change the bicoid protein (controlling the antero-posterior polarity in *Drosophila*) so that it transcriptionally regulates target genes with a specificity like that of the Antp class of protein (Treisman *et al.*, 1989). Yet many proteins that have the same aminoacid at position 9 of the helix 3 or even identical helix 3 have very different effect in embryo development. Thus, the helix 3 differences are not sufficient to account for the specificities of homeodomain action.

The N-terminal is the most divergent region of the homeodomain. The homeodomains of the two *Drosophila* proteins Antp and Scr for example, differ in only five out of sixty residues, and four of these differences are within the N-terminal arm. When these four residues in Antp are replaced by the ones from Scr, the protein behaves like Scr *in vivo* (Zeng *et al.*, 1993). Nonetheless, studies on Ubx-Antp chimeric proteins demonstrated that differences in the N-terminal arm are necessary but not sufficient to change the functions of Ubx into those of Antp (Chan and Mann, 1993).

The region C-terminal to the homeodomain of Ubx has also been reported to influence specificity in experiments in which chimeric Antp-Ubx or Antp-Scr proteins were ectopically expressed in *Drosophila*. However the C-terminal tail does not appear to directly affect monomeric DNA binding (Lin and McGinnis, 1992; Chan and Mann, 1993).

Functional differences among Hox proteins could result from differential interaction with other factors, which modulate their function and/or DNA binding specificity.

In some cases, proteins synergistically activate transcription *in vitro* interacting with other transcription factors or with different components of the general transcriptional machinery. Examples are the POU homeodomain protein Oct-1, which can interact with the strong viral transcription factor VP16

allowing its recruitment to DNA (Lai *et al.*, 1992) and Hunchback and Bicoid, two homeobox-containing proteins involved in the antero-posterior patterning of *Drosophila*, which interact with different subunits of the basal transcription factor TFIID to cooperatively increase its binding to DNA *in vitro* (Sauer *et al.*, 1995).

The differential ability to interact with other proteins requires the presence of different interaction surfaces. Some of them can be represented by residues inside the homeodomain facing away from the DNA which could explain the role of this domain in functional specificity even in the absence of an high specific DNA binding activity.

It has been demonstrated in several systems that other factors interact with homeodomain proteins and refine their target specificities. One example involves the action of the a1 and α2 homeodomain proteins encoded by the budding yeast mating type locus: in haploid α cells, α2 dimers interact with the non-cell-type-specific factor MCM1 to turn off a-specific genes; in diploid cells, the a1 and α2 proteins bind as heterodimers to the promoters of the haploid-specific gene, turning them off (Smith and Johnson, 1992). Thus, DNA binding specificity can also be influenced by heterodimer formation, which can allow a particular protein access to a variety of different binding sites, depending on the partner with which it interacts. Other examples are Phox-1, a homeodomain protein that can bind cooperatively to the c-fos promoter with serum responsive factor, the mammalian homologue of yeast MCM1 (Grueneberg *et al.*, 1992), and HMG1 (High Mobility Group protein 1) which has been shown to interact with the homeodomain of Hox proteins and to enhance their DNA binding and transcriptional activation (Zappavigna et al., 1996). Recently, the *Drosophila* homeoprotein Extradenticle (Exd), and its mammalian counterpart Pbx, have been shown to form heterodimers with several Hox proteins on Hox/Pbx consensus binding sites (Chan *et al.*, 1994; van Dijk and Murre, 1994;

Chang *et al.*, 1995). These proteins are currently considered as functional partners of Hox proteins and their role as Hox cofactors will be discussed separately.

1.4 EXD/PBX PROTEINS AS HOX COFACTORS.

Evidence exists to suggest that homeodomains proteins of a new class, including the *Drosophila* gene product Exd, function as Hox cofactors.

Exd was originally identified as a zygotic, X-linked embryonic lethal mutation that causes homeotic transformations (Wieschaus and Noell, 1986). In null *exd* mutant embryos, some thoracic segments are transformed anteriorly, while some abdominal segments are transformed posteriorly, but the spatial and temporal expression pattern of *Antp*, *Ubx* and *abd-A* expression in these animals are indistinguishable from wild type (Peifer and Wieschaus, 1990). Thus, in an *exd* mutant embryo, at least some Hox genes products are present in the correct place and time, yet they execute incorrect pathways. *Exd* is widely expressed and it has a maternal contribution which generates a uniform distribution of *exd* RNA in the early embryo. This means that the *exd* expression pattern, at least at early developmental stages, does not depend on Hox genes, and maternally-supplied *exd* RNA can fully rescue the segmental defects of the *exd* mutant (Rauskolb *et al.*, 1993). Thus *exd* seems to encode a factor that act in parallel with Hox proteins refining their activities.

Evidence showing that Exd regulates some of the Hox target genes also confirm this hypothesis. For example, the expression pattern of three homeotic target genes in the embryonic visceral mesoderm - *wingless* (*wg*), *tea-shirt* and *decapentaplegic* (*dpp*) - rely on Exd for proper regulation by Abd-A, Antp and Ubx, respectively (Rauskolb and Wieschaus, 1994). In particular, a 303 base pair

regulatory element in the *dpp* promoter was shown to respond to both Ubx and Exd in parasegment 7, implying that the two homeodomain proteins functionally interact on this promoter (Capovilla *et al.*, 1994).

Exd is ubiquitously expressed in early embryonic cells during the period when segmental identities are being determined (Rauskolb *et al.*, 1993) and may act in parallel with many of the Hox family genes as well as with other homeodomain-containing proteins, but it is also expressed in localized regions in the adult fly, where, in addition to being involved in specifying the identity of segments in parallel to what is observed for the embryo, it is also required for functions normally not associated with homeosis, such as the maintenance of the dorsoventral pattern (Gonzales-Crespo and Morata, 1995).

In mammals, three homologs of *exd*, designated Pbx1, Pbx2 and Pbx3 have been identified. Pbx1 was originally identified as a proto-oncogene involved in the t(1;19) chromosomal translocation, which occurs in 20% of paediatric pre-B acute lymphoblastic leukaemia (pre-B ALL). This translocation give rise to E2A-Pbx1, an oncogene encoding a fusion protein containing the N-terminal transcriptional activation domains of the helix-loop-helix transcription factor E2A fused to Pbx1 deleted of a small N terminal region (aminoacids 1 to 89) (Kamps *et al.*, 1990; Nourse *et al.*, 1990). The oncogenic potential of E2A-Pbx1 causes a block in murine myeloid differentiation, and induces myeloid and T-lymphoid leukaemia in transgenic mice. The E2A-Pbx1 induced block in differentiation is specifically dependent on competent DNA binding by the Pbx1 homeodomain, while leukaemia induction in transgenic mice and focus formation in cultured fibroblasts do not require this function (Dedera *et al.*, 1993; Kamps *et al.*, 1996).

Pbx2 and Pbx3 were identified on the basis of their homology to Pbx1 (Monica *et al.*, 1991). Pbx1 and 3 have two differentially spliced transcripts, a

longer form termed "a" and a shorter form termed "b" which have different C-tail regions, while Pbx2 has only an "a" transcript. Pbx1, 2 and 3 are closely related to each other, with a 97% identity at aminoacid level within their homeodomain, but all three proteins diverge significantly near their amino- and carboxy-termini. Their expression is not restricted to particular stages of differentiation or development, as mRNA transcripts of these genes were detected in most fetal and adult tissues and all cell lines with the exception of Pbx1 which is not expressed in the lymphoid lineage.

Another gene homologous to *exd* and Pbx, *ceh-20*, has been also cloned in *C. Elegans*. While all these genes are closely related to each other, they are only distantly related to other homeodomain containing proteins, suggesting that they form a separate family. Together, *exd*, Pbx and *ceh-20* genes are referred to as the PBC family (Bürglin and Ruvkun, 1992). Sequence comparison of members of this family revealed the existence of two extended conserved regions upstream of the homeodomain termed PBC-A and PBC-B (Fig. 1.3). They also contain an atypical homeodomain with an insertion of three aminoacids between helix 1 and helix 2. This insertion is found also in other homeodomain proteins such as yeast MATα-2 and Meis1, which are collectively known as the TALE (three aminoacids loop extension) group (Bürglin, 1997).

The high degree of homology (70%) between PBC genes suggests the conservation throughout evolution of cofactor functions. The genetic data describing a synergy between the homeotic gene products and Exd, have been supported by biochemical data and *in vitro* binding assays: consistent with the proposed role of Hox cofactors, PBC proteins are able to bind cooperatively to DNA with a variety of Hox proteins forming heterodimers with specificities and affinity distinct from those of the individual homeodomain proteins. Using binding sites derived from the *dpp* enhancer as well as synthetic templates, Exd

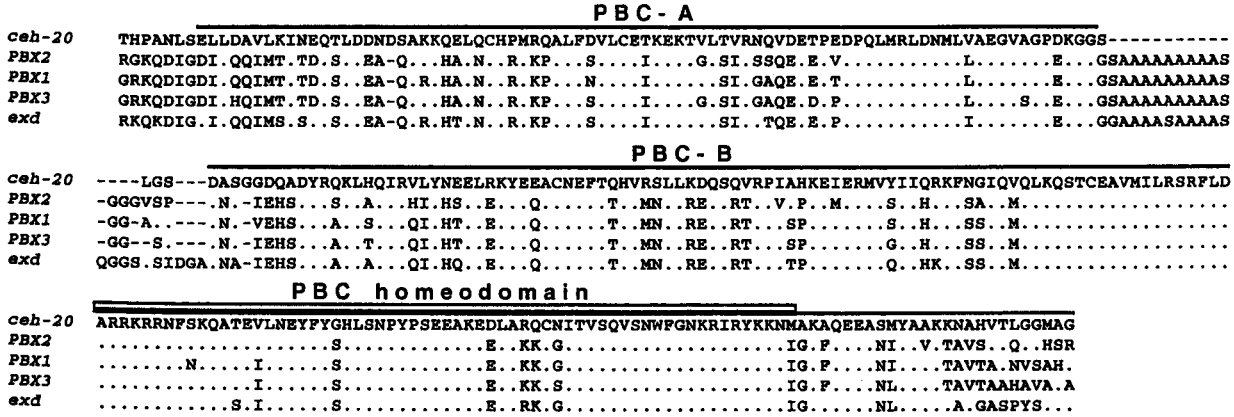


Figure 1.3. Alignment of PBC class genes. The homeodomain region is indicated by a thick bar, and additional regions of sequence conservation are marked by a thin bar. The upstream conserved region has been termed the PBC domain; it can be further subdivided into an A and B region, which are separated in the Pbx genes by a poly-alanine stretch. The PBC homeodomains are atypical, with three extra aminoacids between helix 1 and helix 2.

was demonstrated to cooperatively interact with Ubx and Abd-A (Chan *et al.*, 1994, van Dijk *et al.*, 1994, Sun *et al.*, 1995).

Cooperative binding of Pbx and Hox proteins was first described on the artificial Pbx1 consensus binding site TTGATTGAT, which was derived from selection experiments using degenerated oligonucleotides and recombinant Pbx proteins (Lu *et al.*, 1995). The interaction requires the homeodomains from both the proteins and a highly conserved short stretch of aminoacids located immediately upstream of the Hox homeodomain, known as the YPWM motif (Mavilio *et al.*, 1986). Point mutation experiments demonstrated that the tryptophan (W) and the methionine (M) are the most important residues inside this motif, since their presence is absolutely necessary for complex formation (Chang *et al.*, 1995; Knoepfler and Kamps, 1995; Phelan *et al.*, 1995; Johnson *et al*, 1995).

The formation of an heterodimeric complex with PBC family members has been shown with representative Hox proteins from paralogs 1 to 8, even though the YPWM is variably spaced, 5 to 53 residues, from the N-terminal of the homeodomain. Abd-B and Abd-B-like proteins (paralogs 9 to 13) do not contain a canonical YPWM motif. Their ability to form a complex with Exd/Pbx is controversial: Abd-B-like Hox protein from paralogous groups 9 and 10 have been shown to interact *in vitro* with Pbx1 due to the presence of a W residue 6 aminoacids N-terminal to the homeodomain, but the presence of a W residue located at variable distances upstream the homeodomain in Hox proteins from paralogs 11 to 13 is not sufficient to confer Pbx1 the binding capability to these proteins. In addition, Abd-B itself has been reported to be unable to bind cooperatively to DNA despite the presence of a W residue, also in this case located 6 aminoacids upstream of the homeodomain (van Dijk *et al., 1994*; Chang *et al.*, 1995; Chang *et al.*, 1996; Shen *et al.*, 1997).

Other two homeodomain containing proteins not belonging to the Hox cluster have been shown to form a complex on DNA with Exd/Pbx in a way very similar to Hox proteins: STF-1, which is expressed in the pancreas and the small intestine and also contains a YPWM sequence (Peers *et al.*, 1995), and mammalian an *Drosophila* Engrailed proteins, where a conserved region located in close proximity to the homeodomain contains a W residue whose substitution abolishes cooperative binding on a Pbx consensus site (Peltenburg and Murre, 1996).

Comparing the YPWM sequence in different paralogous groups, only the core (4 aminoacids) is conserved, while members of the same paralogous groups in different species often share a longer region of homology (10-12 aminoacid). Therefore the YPWM motif appears to have coevolved with its associated homeodomain, suggesting that these two domains might function together.

Although direct binding of the YPWM motif to PBC proteins has not yet been demonstrated, several studies using the yeast "two hybrid" assay and DNA binding experiments with mutated proteins suggest that such an interaction exists. For instance, a synthetic peptide containing this sequence can compete for the formation of an Hox/Pbx complex and can stimulate binding by Pbx1 homeodomain alone (Lu and Kamps, 1996b), supporting the hypothesis that the YPWM region could represent a sort of interaction surface for Pbx homeodomain. A sequence immediately C-terminal to Pbx1 homeodomain which is highly conserved in Pbx2 and Pbx3 and predicted to form an α helix, enhances monomeric DNA binding by Pbx1 and also contributes to maximal cooperativity with Hox proteins. Considering all together the data on Pbx/Hox interaction, a structural model has been generated for describing how the Hox and PBC homeodomains bind to DNA (Chan and Mann, 1996; Chang *et al.*, 1996; Lu and Kamps, 1996b; reviewed in

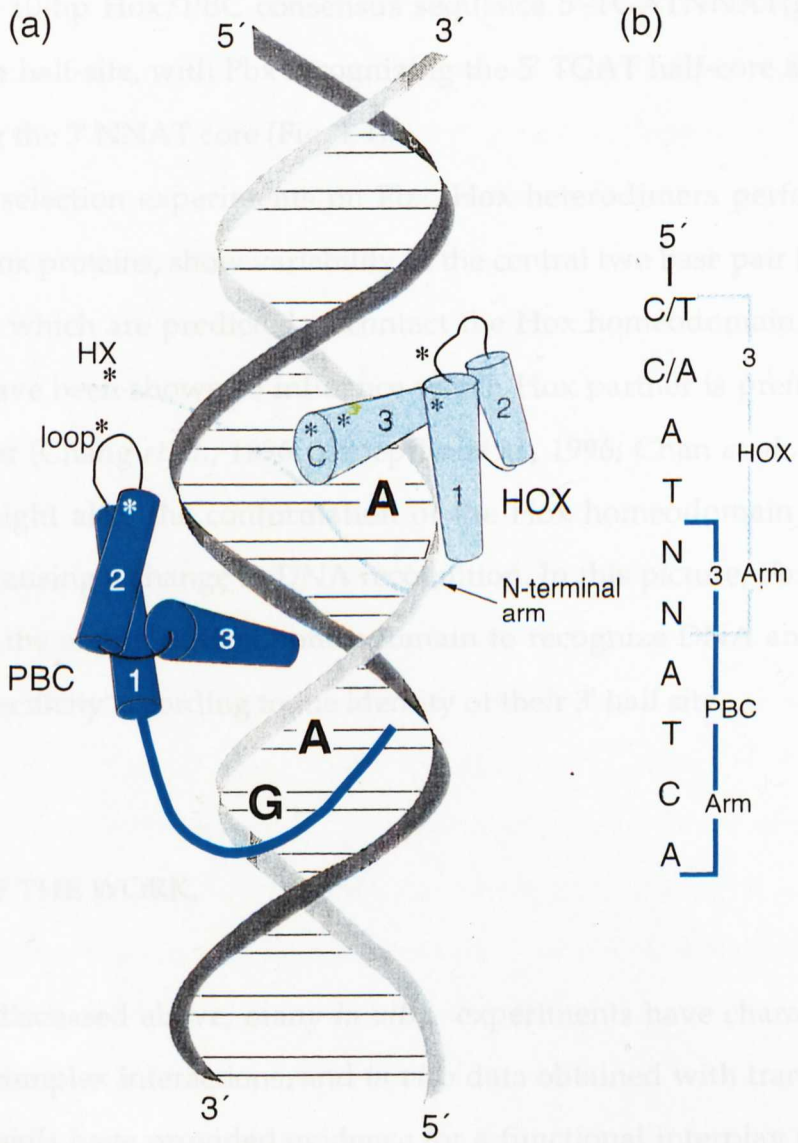


Figure 1.4. A model for the PBC-HOX-DNA complex. **(a)** PBC and HOX homeodomains are shown docked on the DNA. The centre of each binding site, defined by the interaction of the Asn51 (conserved in all homeodomains) and an adenine (**A**), are spaced only by four bp. A guanine (**G**) is predicted to be contacted by Arg55 of the PBC homeodomain in the major groove. The two variable base pairs are contacted by the HOX N-terminal arm, which contribute to specificity, in the minor groove and, potentially, by the PBC third helix in the major groove. The PBC loop (the TALE loop) between helix 1 and 2, the YPWM motif (also called the hexapeptide, HX), the HOX C-tail (C) and individual aminoacids (*) are implicated in the interaction. (Based on known homeodomain-DNA structures). **(b)** Sequence of a HOX-PBC binding site showing the overlapping HOX and PBC half-sites. The choice of HOX protein in the heterodimer is largely due to the sequence of two variable positions (NN) that are contacted by the HOX N-terminal arm. (From Mann and Chan, 1996).

Mann and Chan, 1996). In this model, PBC and Hox proteins bind in tandem on a bipartite, 10-bp Hox/PBC consensus sequence 5'-TGATNNAT(g/t)(g/a)-3' each on one half-site, with Pbx recognizing the 5' TGAT half-core and the Hox recognizing the 3' NNAT core (Fig. 1.4).

Site selection experiments on Pbx/Hox heterodimers performed with different Hox proteins, show variability in the central two base pair (NN) of this 3' half site, which are predicted to contact the Hox homeodomain N-terminal arm and have been shown to influence which Hox partner is preferred in the heterodimer (Chang *et al.*, 1996; Knoepfler *et al.*, 1996; Chan *et al.*, 1997). PBC proteins might alter the conformation of the Hox homeodomain N-terminal arm, thus causing a change in DNA recognition. In this picture, Pbx effectively modulates the ability of Hox homeodomain to recognize DNA and to exhibit binding specificity according to the identity of their 3' half site.

1.5 AIM OF THE WORK.

As discussed above, many *in vitro* experiments have characterized the Hox/Pbx complex interactions, and *in vivo* data obtained with transgenic mice and *Drosophila* have provided evidence for a functional interplay of these two protein families (Pöpperl *et al.*, 1995; Chan *et al.*, 1996; Chan *et al.*, 1997). However, only in very few studies the Hox/Pbx complex has been tested for its ability to cooperatively activate transcription from known target sites, and never without the presence of a heterologous activation domain fused to at least one of the two partners (i.e. either E2A in the E2A-Pbx1 oncogene, or VP16 as fusion with Hox proteins) (Peers *et al.*, 1995; Phelan *et al.*, 1995; Chang *et al.*, 1996).

To directly address the role of Exd/Pbx proteins as modulators of Hox function at the level of transcriptional regulation, we analyzed the transcriptional activity of Hox/Pbx complexes on a well characterized natural Hox target site, the *Hoxb-1* autoregulatory enhancer element b-1ARE, which directs the expression of *Hoxb-1* in the fourth rhombomere (r4) during hindbrain development (Pöpperl *et al.*, 1995).

1.6 SPECIFIC BACKGROUND.

Regional diversity in the vertebrate hindbrain is achieved through a process of segmentation, whereby a series of lineage restricted cellular compartments, termed rhombomeres (r), are formed during early neural development (Fraser *et al.*, 1990; Lumsden and Krumlauf, 1996). These segmental units are correlated with the periodic organization of neurons and the migration of cranial neural crest cells into specific branchial arches, where they differentiate to form distinct skeletal and neurogenic components (Keynes and Krumlauf, 1994; Köntges and Lumsden, 1996). Underlying this cellular organization, the pattern of expression of a number of transcription factors, growth factors, tyrosine kinase receptors and their ligands, have boundaries of expression that are tightly linked to specific hindbrain segments. In particular, Hox genes have anterior boundaries of expression which coincide with specific rhombomere boundaries and which follow the colinearity rule. For instance, the anterior boundaries for member of paralogous groups 2, 3 and 4 lie at rhombomere boundaries r2/3, r4/5 and r6/7, respectively, exhibiting a two-segment repeat pattern (Fig. 1.5). Members of paralogous group 1 are the only exception to this colinearity, since *Hoxb-1* and *Hoxa-1* have a boundary at r3/4, which is posterior to the

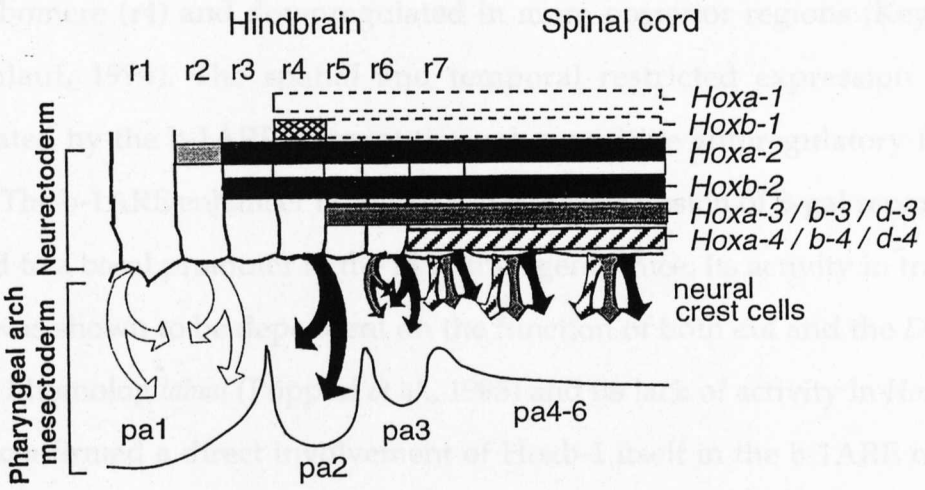


Figure 1.5. *Hox* genes expression domains in the rhombomeres and pharyngeal arches of the mouse embryo. Individual rhombomeres are drawn, and the arrows represent the specific migratory pathway of neural crest cells into pharyngeal arches. Note the two-segments periodicity between the rostral expression boundaries of genes from homology groups 2, 3 and 4 (with the exception of *Hoxa-2* which is also weakly expressed in rhombomere 2), the absence of *Hox* transcripts in the first arch mesectoderm and the increasing combinations of *Hox* transcripts towards posterior pharyngeal arches. The expression domains of *Hoxa-1* and *Hoxb-1* (dashed lines) are seen at earlier stages, prior to the formation of rhombomere boundaries, and only *Hoxb-1* expression is maintained later in rhombomere 4. (From Favier and Dollé, 1997).

expression domain of group 2 genes. Another member of this group, *Hoxd-1*, is not expressed in the central nervous system (Keynes and Krumlauf, 1994; Lumsden and Krumlauf, 1996; Maconochie *et al.*, 1996).

In the early mouse embryo (8.0 dpc), *Hoxb-1* has an anterior expression limit in the hindbrain mapping at the boundary of rhombomeres 3 and 4 (r3/4). In later stages (9.5 dpc), *Hoxb-1* expression is upregulated in the fourth rhombomere (r4) and downregulated in more posterior regions (Keynes and Krumlauf, 1994). The spatial and temporal restricted expression in r4 is mediated by the b-1ARE element through a positive autoregulatory feedback loop. The b-1ARE enhancer is able to drive the expression of β-gal reporter gene linked to a basal promoter in the r4 of transgenic mice. Its activity in transgenic flies was shown to be dependent on the function of both *exd* and the *Drosophila Hoxb-1* homolog *labial* (Pöpperl *et al.*, 1995) and its lack of activity in *Hoxb-1* null mice confirmed a direct involvement of Hoxb-1 itself in the b-1ARE mediated regulation in r4 (Studer *et al.*, 1996).

The 148-bp b-1ARE enhancer is located in the 5' flanking region of the mouse *Hoxb-1* gene and contains three related sequence motifs, namely repeats 1, 2 and 3 (R1, R2, R3), which are nearly identical to the Pbx consensus binding site defined *in vitro* (Fig. 1.6). These repeats are necessary for b-1ARE activity, since mutations of all three of them abolish the function of the enhancer. Repeat 1 and repeat 3 have been shown to mediate cooperative binding *in vitro* of *Hoxb-1* and Exd proteins. However, repeat 3 in particular seems to be the most crucial site for the activity and three copies of this repeat cloned in front of a basal promoter are sufficient to direct r4 restricted expression in mice (Pöpperl *et al.*, 1995).

Sequence comparison in mouse, chicken and pufferfish revealed that the b-1ARE element is very conserved inside a wider and more divergent upstream region of the *Hoxb-1* gene. Moreover, the pufferfish sequence is also able to

generate r4 restricted expression of a reporter construct in transgenic mice, demonstrating a functional conservation of this element in different vertebrate species (Pöpperl *et al.*, 1995).

TTCTTTCCAAAAAGTCTTTGAAGAAAGATG**T**TTTGACGCTTCCATGTCG 50

CTCTC**AGATGGAT**GGGCTCAGAG**TGATTGAA**GTGTCTTTGTCATGCTAAT 100

GATTGGGGGG**TGATGGAT**GGGCGCTGGGACTGCCAAACTC 140

R1	**AGATGGAT**
R2	**TGATTGAA**
R3	**TGATGGAT**
pbx consensus	**TGATTGAT**

Figure 1.6. Sequence of the mouse b-1ARE r4 enhancer. The three Pbx consensus repeats (R1, R2 and R3) are in bold. Another region which is also conserved in the Hoxb-1 r4 enhancer of different vertebrate species is underlined. It contains a TAAT motif in the context of an octamer-binding protein consensus site, which is marked by a thick bar.

Chapter 2

RESULTS

2.1 HOXB1 AND PBX1 COOPERATIVELY ACTIVATE TRANSCRIPTION THROUGH THE *HOXB-1* GENE AUTOREGULATORY ELEMENT.

The transcriptional activity of the HOXB1/Pbx1 complex was assayed in transient co-transfection experiments, using the *Hoxb-1* r4 autoregulatory enhancer (b-1ARE) (Pöpperl *et al.*, 1995) as a responsive element. The pAdMLARE reporter construct contains a 148-bp fragment from the b-1ARE upstream of the Adenovirus major late (AdML) minimal promoter in the pAdMLluc luciferase reporter gene (Fig. 2.1A). Transfection of COS7 cells with the pAdMLARE reporter (b-1ARE in Fig. 2.1B) together with a construct expressing either HOXB1 or Pbx1 under the control of the SV40 promoter (pSGHOXB1 and pSGPbx1 respectively) did not significantly stimulate the reporter basal activity. Conversely, co-transfection of both the Pbx1 and the HOXB1 expressors caused a 9-fold enhancement of the pAdMLARE activity but not of the enhancer-less pAdMLluc reporter transfected as a control (Fig. 2.1B).

Among the three related sequence motifs present in the b-1ARE, the 3'-most repeat 3 (R3) was shown to mediate most of the activity of the b-1ARE in transgenic mice (Pöpperl *et al.*, 1995). To study the relative role of this motif in mediating transcriptional activation from the HOXB1/Pbx1 complex, a knockout mutation of R3 within the b-1ARE reporter (pAdMLAREmR3, Fig. 1A) was generated. Co-transfection of pSGHOXB1 and pSGPbx1 expressors

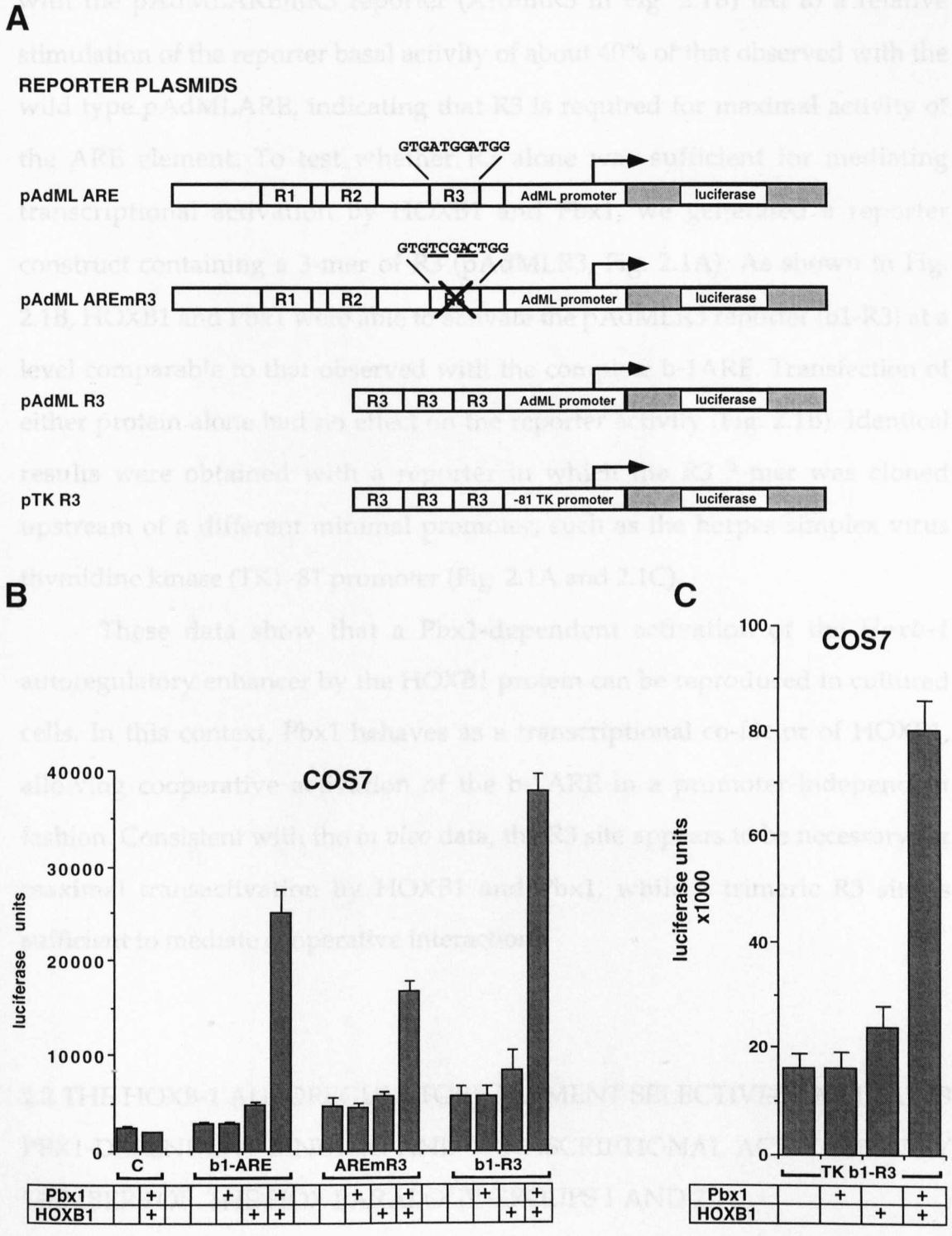


Figure 2.1. **(A)** Schematic representation of the reporter constructs used in transfection assays. R1, R2 and R3 indicate the evolutionarily conserved repeated sequences within the b-1ARE genomic region. Mutated bases within the R3 repeat are underlined. A grey box represents the luciferase reporter gene. **(B, C)** Luciferase activity, in arbitrary units, assayed from extracts of transiently transfected COS7 cell line. Cells were transfected with 4 µg of the SV40-driven HOXB1, and/or Pbx1 expression constructs, together with 8 µg of the indicated reporter constructs (C= pAdMLluc control reporter). 0.2 µg of the pCMVß-gal plasmid were cotransfected in all experiments as an internal standard. Bars represent the mean ± standard error of the mean (S.E.M.) of at least four independent experiments.

with the pAdMLAREmR3 reporter (AREmR3 in Fig. 2.1B) led to a relative stimulation of the reporter basal activity of about 40% of that observed with the wild type pAdMLARE, indicating that R3 is required for maximal activity of the ARE element. To test whether R3 alone was sufficient for mediating transcriptional activation by HOXB1 and Pbx1, we generated a reporter construct containing a 3-mer of R3 (pAdMLR3, Fig. 2.1A). As shown in Fig. 2.1B, HOXB1 and Pbx1 were able to activate the pAdMLR3 reporter (b1-R3) at a level comparable to that observed with the complete b-1ARE. Transfection of either protein alone had no effect on the reporter activity (Fig. 2.1B). Identical results were obtained with a reporter in which the R3 3-mer was cloned upstream of a different minimal promoter, such as the herpes simplex virus thymidine kinase (TK) -81 promoter (Fig. 2.1A and 2.1C).

These data show that a Pbx1-dependent activation of the *Hoxb-1* autoregulatory enhancer by the HOXB1 protein can be reproduced in cultured cells. In this context, Pbx1 behaves as a transcriptional co-factor of HOXB1, allowing cooperative activation of the b-1ARE in a promoter-independent fashion. Consistent with the *in vivo* data, the R3 site appears to be necessary for maximal transactivation by HOXB1 and Pbx1, while a trimeric R3 site is sufficient to mediate cooperative interaction.

2.2 THE HOXB-1 AUTOREGULATORY ELEMENT SELECTIVELY MEDIATES PBX1-DEPENDENT BINDING AND TRANSCRIPTIONAL ACTIVATION BY MEMBERS OF THE HOX PARALOGY GROUPS 1 AND 2.

We next tested whether other HOX gene products could transactivate the *Hoxb-1* ARE in combination with Pbx1. Twelve HOX genes belonging to different clusters and/or paralogy groups were co-expressed with Pbx1 in

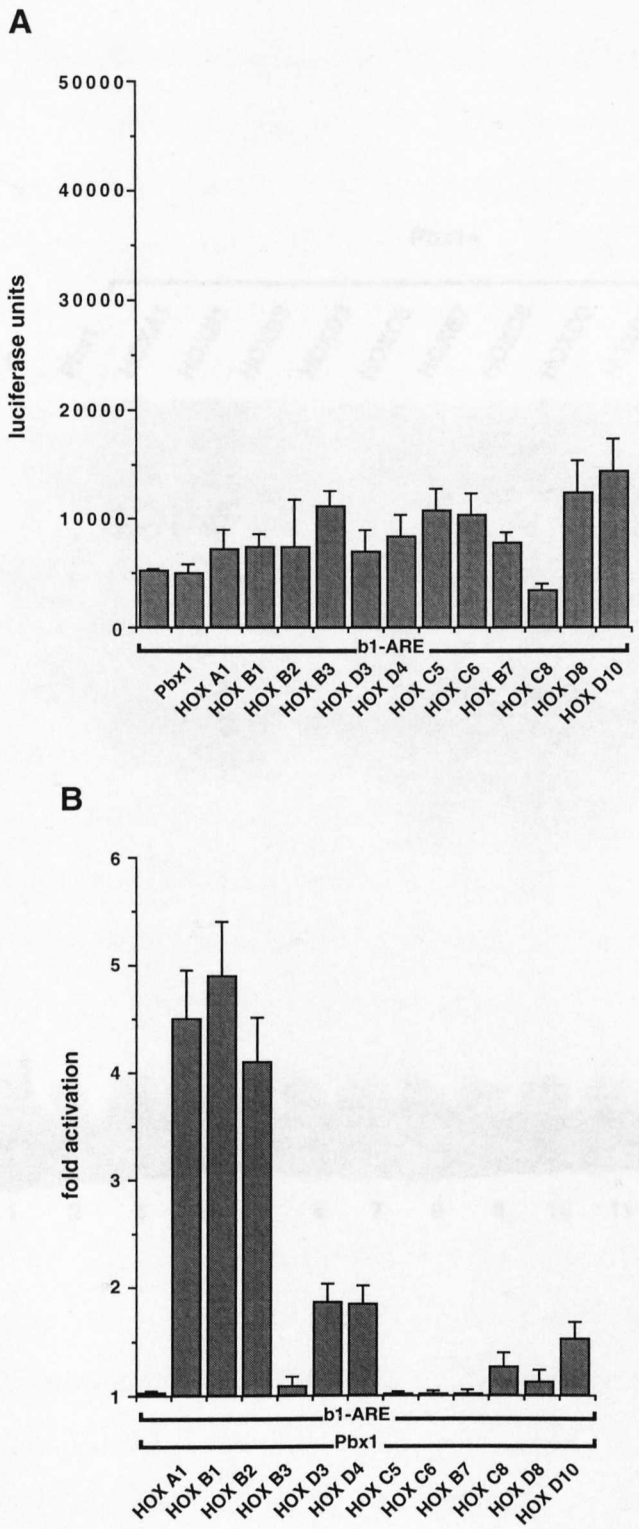


Figure 2.2. Luciferase activity from transiently transfected Cos7 cells. **(A)** Cells were transfected with 4 μg of SV40-driven expression vectors producing the indicated HOX full length proteins, and 8 μg of the pAdMLARE reporter construct. Bars represent mean luciferase units ± S.E.M. of at least four independent experiments. **(B)** Cells were transfected with 4 μg of the same HOX expressors as in **(A)** together with 8 μg of the Pbx1 expression vector, and 8 μg of the pAdMLARE reporter construct. Bars represent mean fold activation over the activity obtained transfecting the HOX proteins alone with pAdMLARE ± S.E.M. of at least four independent experiments.

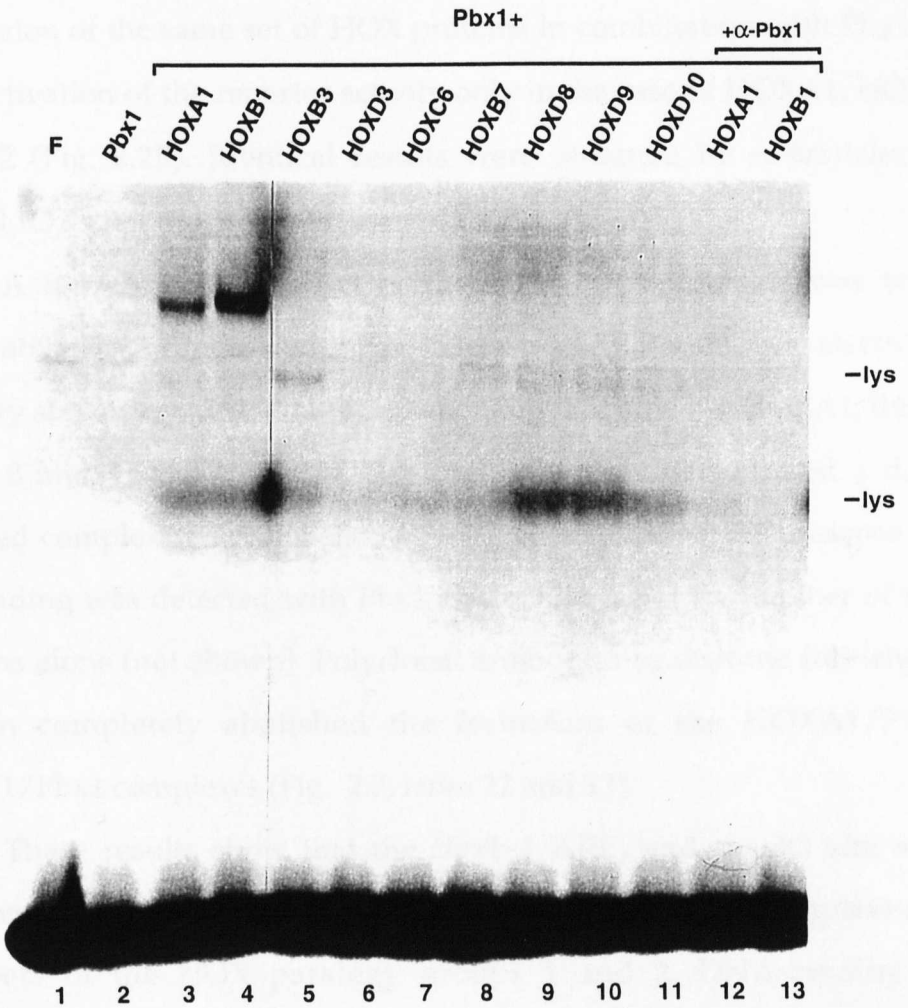


Figure 2.3. Electrophoretic mobility shift assay (EMSA) of a 30-bp double-stranded oligonucleotide representing the R3 site of the b-1ARE. The Pbx1 and the indicated HOX proteins were produced and labelled by coupled transcription/translation from their expression vectors in rabbit reticulocyte lysates. Lysates (1-4 µl depending on protein content) were mixed with the radiolabelled R3 oligonucleotide in a binding reaction, and subjected to EMSA. (F) Free probe. Anti-Pbx1 rabbit antiserum (α-Pbx1), was added to the binding reactions containing HOXA1 and HOXB1 (lanes 12 and 13). (lys), complexes arising from endogenous reticulocyte lysate binding activities.

transient transfections of COS7 cells, and tested for cooperative transcriptional activation of the pAdMLARE reporter. As shown in Fig. 2.2A, none of the tested HOX proteins was able to transactivate alone the pAdMLARE reporter. Expression of the same set of HOX proteins in combination with Pbx1 led to a transactivation of the reporter activity only in the case of HOXA1, HOXB1 and HOXB2 (Fig. 2.2B). Identical results were obtained by co-transfecting the pAdMLR3, containing the R3 trimer (data not shown).

A representative subset of the same HOX proteins was tested for cooperative DNA-binding with Pbx1 to the b-1ARE R3 site in an electrophoretic mobility shift assay (EMSA). As shown in Fig. 2.3, only the HOXA1, the HOXB1 (lanes 3 and 4) and the HOXB2 proteins (not shown) formed a detectable retarded complex with a labelled R3 oligonucleotide in the presence of Pbx1. No binding was detected with Pbx1 alone (lane 2), or with either of the HOX proteins alone (not shown). Polyclonal antibodies against the full-length Pbx1 protein completely abolished the formation of the HOXA1/Pbx1 and HOXB1/Pbx1 complexes (Fig. 2.3, lanes 12 and 13).

These results show that the *Hoxb-1* ARE, and the R3 site within it, selectively mediate transcriptional activation by HOX/Pbx1 complexes made of members of the HOX paralogy groups 1 and 2. DNA-binding *in vitro* experiments suggest that this selectivity relies on the differential recognition of the R3 site by HOX proteins, allowing only members of paralogy groups 1 and 2 to bind this sequence cooperatively with Pbx1. Interestingly, in addition to *Hoxb-1*, both *Hoxa-1* and *Hoxb-2* are expressed in rhombomere 4 during development (reviewed in Keynes and Krumlauf, 1994).

2.3 SPECIFIC TARGET RECOGNITION BY THE HOXB1/PBX1 COMPLEX IS ENCODED IN THE HOX HOMEODOMAIN N-TERMINAL ARM, AND CAN BE TRANSFERRED TO HOXB3 BY DOMAIN SWAPPING.

To identify the HOX protein domain responsible for the differential activation of the R3 site *in vivo*, we carried out a domain-swap experiment between HOXB1 and another HOX protein unable to activate transcription through the b-1ARE in combination with Pbx1, i.e., HOXB3. As discussed in Chapter 1, several studies had indicated that specificity of action of HOX genes resides mostly in the homeodomain and its flanking regions (reviewed in Krumlauf, 1994). In particular, the flexible N-terminal arm of the homeodomain was shown to be crucial for HOX specificity both *in vivo* (Furukubo-Tokunaga *et al.*, 1993; Lin and McGinnis, 1992; Zeng *et al.*, 1993) and in transfected cells (Zappavigna *et al.*, 1994). The conserved "YPWM" motif was identified as an important interaction surface between Hox and Pbx proteins (reviewed in Mann, 1996). Therefore, we decided to replace a region encompassing the "YPWM" motif (FDWM in HOXB1 and FPWM in HOXB3) and the homeodomain N-terminus in the HOXB3 protein (aa 130 to 201) with the corresponding region of HOXB1 (aa 175 to 211) (Fig. 2.4A), and to assay the resulting chimeric protein (HOXB3/B1HN in Fig. 2.4A) for transcriptional activation of the pAdMLARE reporter in combination with Pbx1. As shown in Fig 2.4B, while the wild-type HOXB3 was unable to transactivate the pAdMLARE reporter, either alone or in combination with Pbx1, the HOXB3/B1HN chimera could activate the reporter together with Pbx1 to a level comparable to that obtained with HOXB1. Accordingly, the HOXB3/B1HN chimera could cooperatively bind to the R3 site in EMSA with Pbx1 (Fig. 2.5).

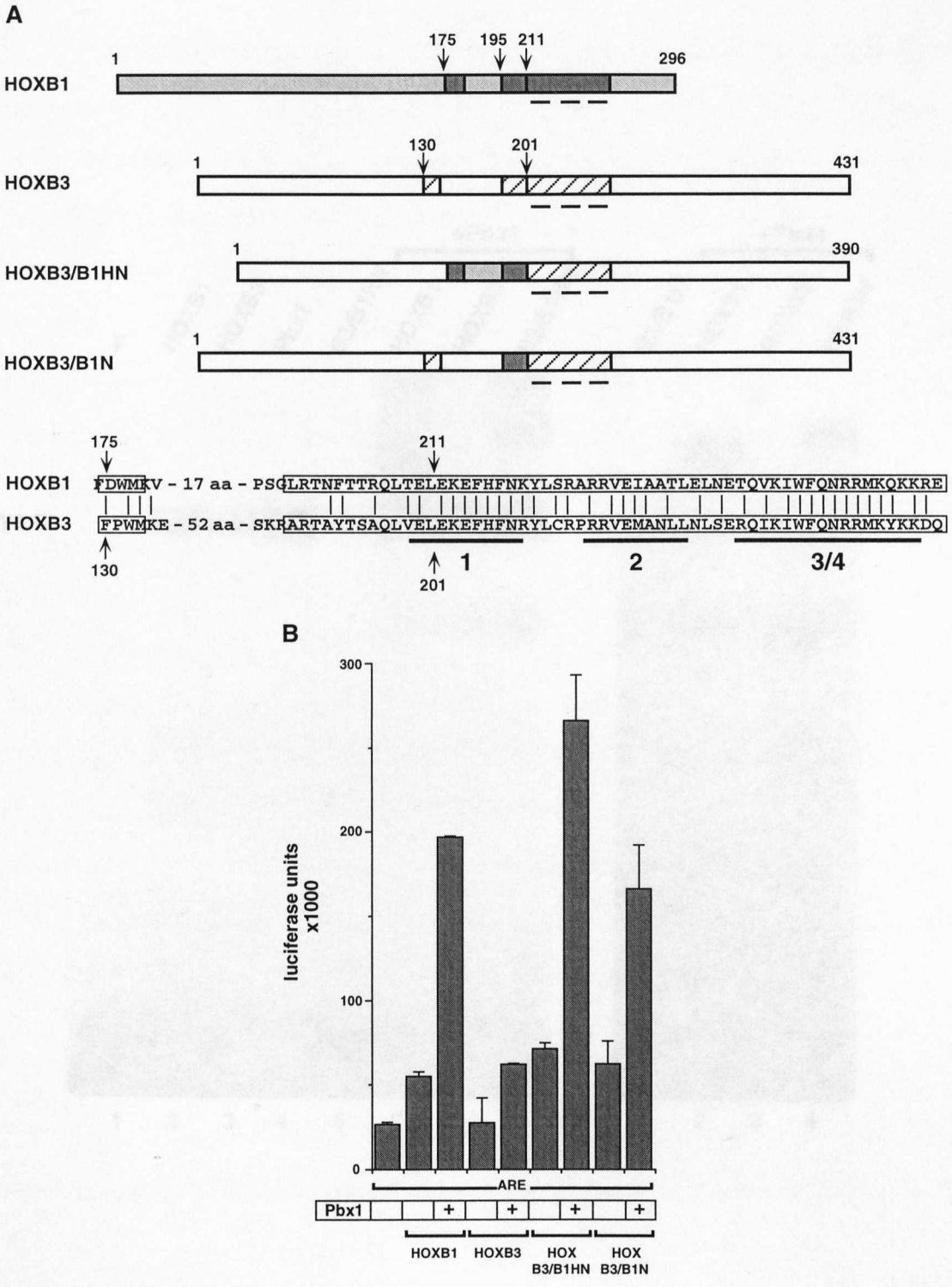


Figure 2.4. **(A)** Schematic representation of the HOXB1, the HOXB3 and the HOXB3/B1 chimeric proteins. Numbers indicate amino acid positions, dark grey or striped boxes indicate the homeodomain, horizontal lines represent the helical regions within the homeodomain. The amino acid sequences of the HOXB1 and HOXB3 homeodomains are shown aligned, vertical bars highlight conserved residues, boxed sequences represent the "YPWM" and the homeodomain regions respectively. **(B)** Luciferase activity from transiently transfected COS7. Cells were transfected with 8 µg of the pAdMLARE reporter (b-1ARE), together with 8 µg of the Pbx1 expressor where indicated, and with 4 µg of the HOXB1, the HOXB3, and the HOXB3/B1 chimera expressors where indicated. 0.2 µg of the pCMVß-gal plasmid were cotransfected in all experiments as an internal standard. Bars represent the mean luciferase activity ± S.E.M. of at least three independent experiments.

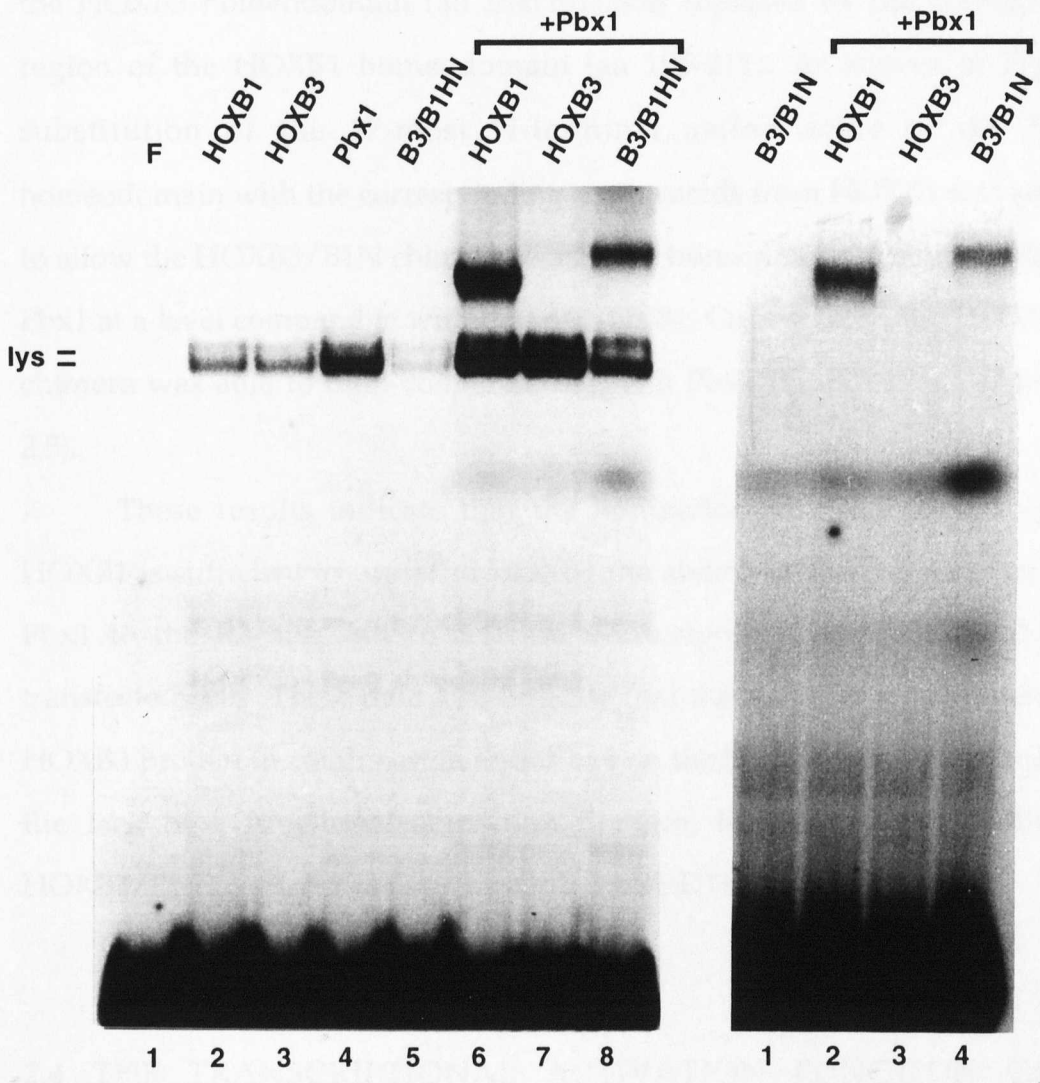


Figure 2.5. EMSA of the 30-bp double-stranded R3 oligonucleotide. The Pbx1, HOXB1, HOXB3, and chimeric HOXB3/B1 proteins were produced and labelled by coupled transcription/translation from their expression vectors in rabbit reticulocyte lysates. Lysates (1-4 µl) were mixed with the radiolabelled R3 oligonucleotide in a binding reaction, and subjected to EMSA. (F) Free probe. (lys), complexes arising from endogenous reticulocyte lysate binding activities.

To further narrow down the region responsible for the differential activity of HOXB1 and HOXB3 on the R3 site, we constructed an additional chimeric mutant (HOXB3/B1N in Fig. 2.4A), in which the N-terminal portion of the HOXB3 homeodomain (aa 184-201) was replaced by the corresponding region of the HOXB1 homeodomain (aa 195-211). As shown in Fig. 2.4B, substitution of the 17-most N-terminal amino acids of the HOXB3 homeodomain with the corresponding amino acids from HOXB1 was sufficient to allow the HOXB3/B1N chimera to activate transcription in combination with Pbx1 at a level comparable with that of HOXB1. Consistently, the HOXB3/B1N chimera was able to bind cooperatively with Pbx1 the R3 site in EMSA (Fig. 2.5).

These results indicate that the homeodomain N-terminus region of HOXB1 is sufficient to confer to HOXB3 the ability to bind cooperatively with Pbx1 to the R3 site and to activate transcription through the b-1ARE in transfected cells. These data also indicate that the absence of activation by the HOXB3 protein in combination with Pbx1 on the b-1ARE element is not due to the lack of a functional activation domain, but rather to inability of a HOXB3/Pbx1 complex to form on the b-1ARE R3 site.

2.4 THE TRANSCRIPTIONAL ACTIVATION FUNCTION OF THE HOXB1/PBX1 COMPLEX RESIDES MAINLY IN THE HOXB1 ACTIVATION DOMAIN.

We took advantage of the activity of HOXB1 and Pbx1 on the *Hoxb-1* ARE to dissect the functional properties of the HOXB1/Pbx1 complex in transcriptional activation. To identify the protein domain(s) responsible for transcriptional activation, we generated a series of deletion mutants in the

HOXB1 and Pbx1 proteins (Fig. 2.6A), and tested them for the ability to activate the pAdMLR3 reporter in transfected COS7 cells. Deletion of the N-terminal region (amino acids 1-155) of HOXB1 (HOXB1HD in Fig. 2.6A) caused a reduction in the activity of the HOXB1/Pbx1 complex of about 70% with respect to the complex containing the wild type HOXB1 (Fig. 2.6B). The HOXB1HD mutant was still able to bind cooperatively with Pbx1 on the b-1ARE R3 site (Fig. 2.7, lane 8). Conversely, two N-terminal deletions of Pbx1 (Pbx1aΔ1-140 and Pbx1aΔ1-230 in Fig. 2.6A), lacking one or both of the conserved PBC-A and PBC-B domains previously identified in *exd*/Pbx family members (Burglin and Ruvkun, 1992), led to cooperative activation of the pAdMLR3 reporter activity at levels comparable to, or higher than, those observed with the wild type Pbx1 (Fig. 2.6B). Interestingly, the Pbx1aΔ1-230 mutant was able to activate the pAdMLR3 reporter even in absence of HOXB1 (Fig. 2.6B). These data indicate that the transcriptional activating function of the HOXB1/Pbx1 complex on the b-1ARE element resides essentially in the HOXB1 N-terminal region. This region contains a *bona fide* activation domain, since it was active also in the context of a fusion with the yeast GAL4 DNA binding domain (aa 1-147) on a GAL4-responsive reporter in transfected COS7 cells (results not shown).

To test whether binding of either HOXB1 or Pbx1 to DNA is necessary for activation of the pAdMLR3 reporter, we generated and tested DNA-binding defective mutants in both proteins. A HOXB1 mutant lacking helix 3 of the homeodomain (B1Δ236-274 in Fig. 2.6A) did not stimulate the pAdMLR3 reporter basal activity when expressed in combination with Pbx1 (Fig. 2.6C). Similarly, a Pbx1 deletion mutant lacking three aminoacids within helix 3 of the homeodomain (Pbx1aΔ283-285, in Fig. 2.6A) showed no activity when coexpressed with HOXB1 (Fig. 2.6C). When tested in EMSA, neither B1Δ236-274 (not shown) nor Pbx1aΔ283-285 (Fig. 2.7, lane 12) were able to bind

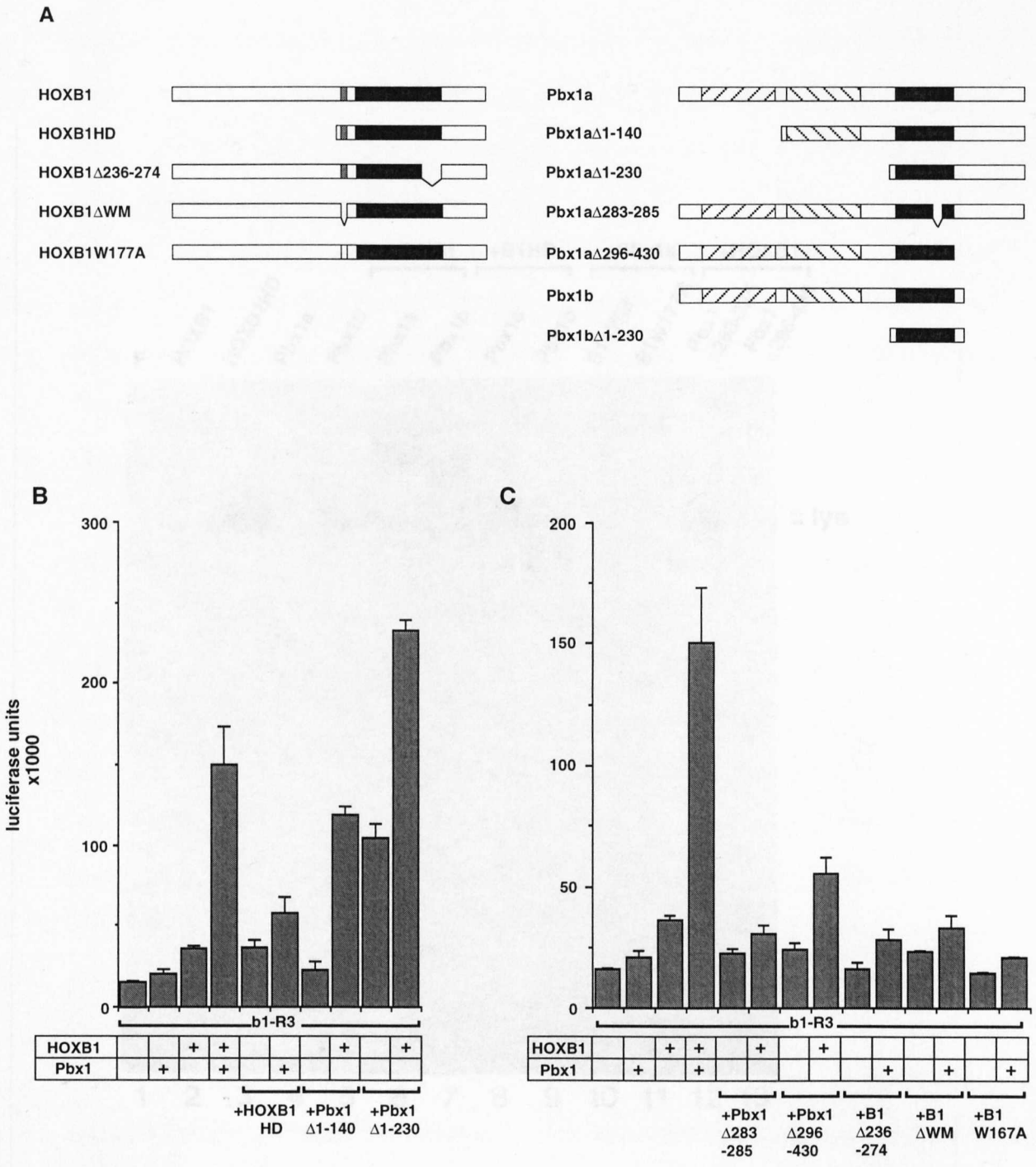


Figure 2.6. Mutational analysis of the HOXB1/Pbx1 complex transcriptional activation. **(A)** Schematic representation of the HOXB1 (left side) and of the Pbx1 (right side) mutants. Black boxes represent the homeodomains, grey boxes the FDWM, a white box the W177A substitution within the motif, and striped boxes the PBC-A and PBC-B domains conserved in exd/Pbx proteins. **(B, C)** Luciferase activity from transiently transfected COS7. Cells were transfected with 8 μg of the pAdMLR3 reporter (b1-R3), together with 8 μg of expressor plasmids for Pbx1 and its mutants, and with 4 μg of expressor plasmids for HOXB1 and its mutants in the indicated combinations. Mutants are identified in **(A)**. 0.2 μg of the pCMVß-gal plasmid were cotransfected in all experiments as an internal standard. Bars represent the mean luciferase activity ± S.E.M. of at least three independent experiments.

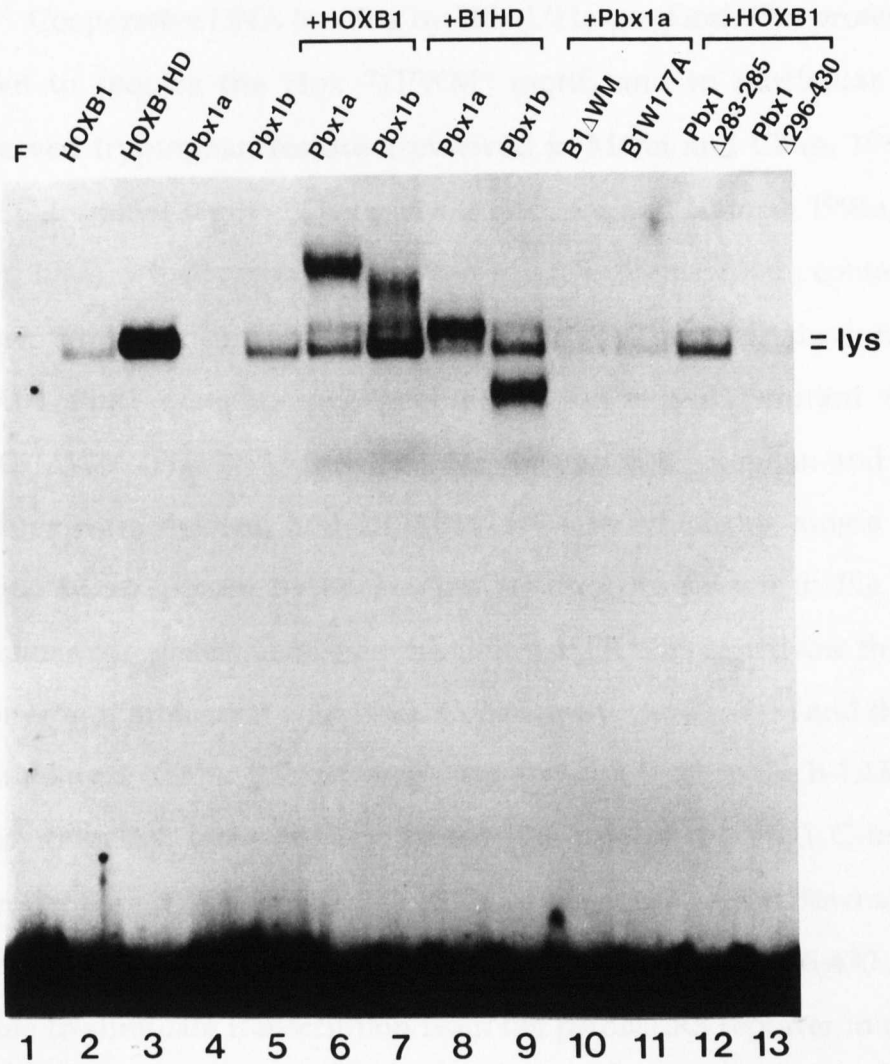


Figure 2.7. Cooperative DNA-binding of HOXB1 and Pbx1 mutants to the 30-bp double-stranded oligonucleotide representing the R3 site of the b-1ARE. Pbx1, HOXB1, and their mutants, were produced and labelled by coupled transcription/translation from their corresponding expression vectors in rabbit reticulocyte lysates. Lysates (1-4 µl) were mixed with the radiolabelled R3 oligonucleotide in a binding reaction, and subjected to EMSA. (F) Free probe. (lys), complexes arising from endogenous reticulocyte lysate binding activities.

cooperatively to the R3 site in association with the wild-type Pbx1 or HOXB1 respectively. Thus, mutations affecting DNA binding of either HOXB1 or Pbx1 protein disrupt the formation of a complex *in vitro* and suppress transcriptional activation in transfected cells.

Cooperative DNA binding by HOM/Hox and *exd*/Pbx proteins has been shown to require the Hox "YPWM" motif, and in particular the highly conserved tryptophan residue (reviewed in Mann and Chan, 1996), and the Pbx1 C-terminal region (Chang *et al.*, 1995; Lu and Kamps, 1996a; Mann and Chan, 1996), which are supposed to mediate protein-protein contacts between the two proteins. To test the role of the "YPWM" motif in the function of the HOXB1/Pbx1 complex, we generated two HOXB1 mutant derivatives, HOXB1ΔWM (Fig. 2.6A) in which the conserved tryptophan and methionine residues were deleted, and HOXB1W177A in which the single tryptophan residue was replaced by an alanine residue. As shown in Fig. 2.6C, both mutations completely abolished the ability of HOXB1 to activate the pAdMLR3 reporter in combination with Pbx1. Consistently, the B1ΔWM and the B1W177A mutants were unable to bind cooperatively with Pbx1 to the b-1ARE R3 site in EMSA (Fig. 2.7, lanes 10, 11). To test the role of the Pbx1 C-terminus, we generated a mutant carrying a deletion of the entire region downstream to the homeodomain (Pbx1aΔ296-430 in Fig. 2.6A). The Pbx1aΔ296-430 mutant was unable to stimulate transcription from the pAdMLR3 reporter in combination with HOXB1 (Fig. 2.6C). Consistently, this mutant did not form a complex with HOXB1 on the R3 site in EMSA (Fig. 2.7, lanes 12, 13). Production and nuclear localisation of the mutants was tested by Western blot analysis of nuclear extracts from transfected cells (not shown).

Taken together, these results indicate that the DNA-binding domains of both HOXB1 and Pbx1, and the putative protein-protein interaction domains on

both proteins, are all individually necessary for the assembly of a cooperative DNA-binding and transcriptionally active complex.

2.5 THE C-TERMINUS OF THE PBX1a ISOFORM CONTAINS AN ACTIVATION DOMAIN WHICH IS ABSENT IN THE PBX1b SPLICING VARIANT.

As reported above, the deletion of the N-terminal region of HOXB1 caused a significant but not complete reduction of the transcriptional activity of the HOXB1/Pbx1 complex (HOXB1HD mutant in Fig. 2.6B). This observation led us to assume that the Pbx1 protein could contribute in part to the transcriptional activity of the HOX/Pbx complex, and therefore be responsible for the residual activity of the B1HD/Pbx1 complex on the b-1ARE element. To identify a potential activation domain within the Pbx1 protein, we focused on the C-terminal region, since deletions within the N-terminus did not reduce the activity of the HOXB1/Pbx1 complex (Δ1-140 and Δ 1-230 mutants in Fig. 2.6B). In particular, we tested the activity of the Pbx1b isoform, a naturally occurring, alternative splicing variant of the full-length Pbx1 protein (or Pbx1a), which is 83 amino acids shorter at the C-terminus (Monica *et al.*, 1991). As shown in Fig. 2.8A, co-expression of Pbx1b and HOXB1 led to activation of the pAdMLR3 reporter at levels of ~ 50% of those obtained with Pbx1a, while co-expression of Pbx1b and the HOXB1 mutant carrying a deletion of the N-terminal, major activation domain (HOXB1HD) showed no residual activity. These data indicate that the C-terminal 83 amino acids of the Pbx1a isoform carry an activation domain which is not contained in Pbx1b, and which contributes to the overall transcriptional activity of the HOXB1/Pbx1 complex.

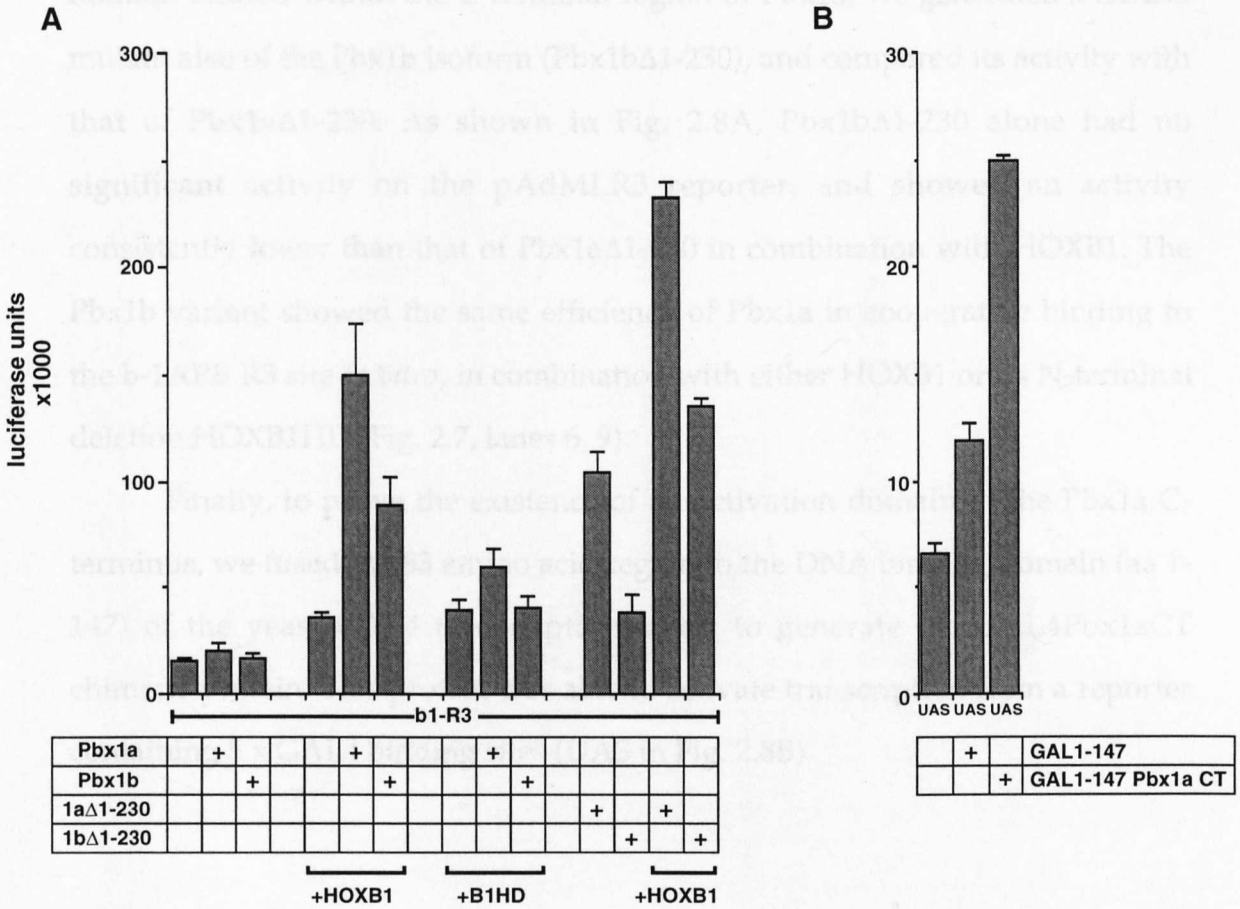


Figure 2.8. **(A)** Luciferase activity from transiently transfected COS7. Cells were transfected with 8 µg of the pAdMLR3 reporter (b1-R3), together with 8 µg of the Pbx1a, Pbx1b, Pbx1aΔ1-230 (1aΔ1-230) and Pbx1bΔ1-230 (1bΔ1-230) expressors where indicated, and with 4 µg of the HOXB1, or HOXB1HD (B1HD) expressors where indicated. **(B)** Cells were transfected with 8 µg of the pTUAS reporter (UAS) together with 4 µg of the pGAL1-147 (GAL1-147) or the pGAL1-147Pbx1aCT (GAL1-147Pbx1aCT) expression construct. 0.2 µg of the pCMVß-gal plasmid were cotransfected in all experiments as an internal standard. Bars represent the mean luciferase activity ± S.E.M. of at least five independent experiments.

As described above, a Pbx1a deletion mutant lacking amino acids 1-230 (Pbx1aΔ1-230) was able to activate the pAdMLR3 reporter even in the absence of HOXB1 (Fig. 2.6B). To test whether this activity relied on the activation domain located within the C-terminal region of Pbx1a, we generated a Δ1-230 mutant also of the Pbx1b isoform (Pbx1bΔ1-230), and compared its activity with that of Pbx1aΔ1-230. As shown in Fig. 2.8A, Pbx1bΔ1-230 alone had no significant activity on the pAdMLR3 reporter, and showed an activity consistently lower than that of Pbx1aΔ1-230 in combination with HOXB1. The Pbx1b variant showed the same efficiency of Pbx1a in cooperative binding to the b-1ARE R3 site *in vitro*, in combination with either HOXB1 or its N-terminal deletion HOXB1HD (Fig. 2.7, lanes 6, 9).

Finally, to prove the existence of an activation domain in the Pbx1a C-terminus, we fused the 83 amino acid region to the DNA binding domain (aa 1-147) of the yeast GAL4 transcription factor to generate the GAL4Pbx1aCT chimeric protein. This protein was able to activate transcription from a reporter containing 5 x GAL4 binding sites (UAS in Fig. 2.8B).

2.6 THE TRANSCRIPTIONAL ACTIVITY OF THE HOX/PBX COMPLEX IS INCREASED IN EMBRYONAL CARCINOMA CELL LINES.

P19 is a murine embryonal carcinoma cell line with the potential to differentiate towards different cell types, most notably cells of the neuroectodermal lineages. Expression constructs for HOXA1, HOXB1 and HOXB2, which showed cooperative binding on the R3 site and transcpritional activity together with Pbx1 both on the b-1ARE element and on the trimeric R3 site (Fig. 2.2 and not shown), were co-transfected also in P19 cells, either with the pAdMLARE or with pAdMLR3 reporter constructs (Fig. 2.9). Contrarily to

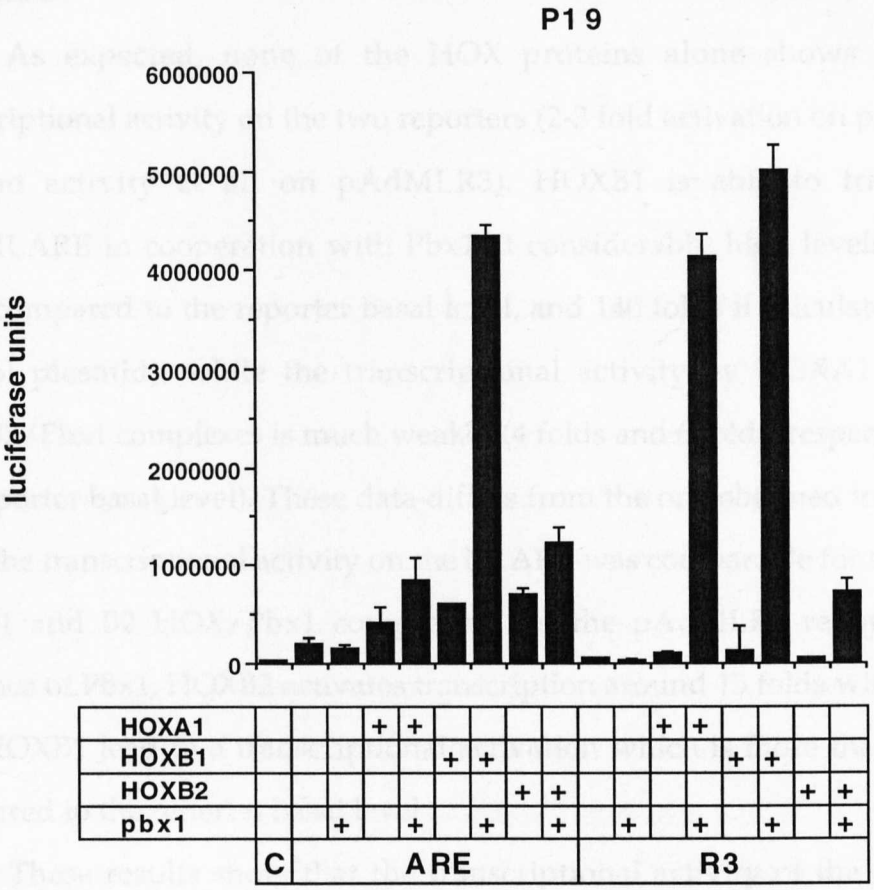


Figure 2.9. Luciferase activity from transiently transfected P19 cells. 8 µg of the pAdMLARE (ARE) or pAdMLR3 reporters (R3) were cotransfected together with 8 µg of the Pbx1a and 4 µg of HOXA1 HOXB1 and HOXB2 expressors where indicated. 0.2 µg of the pCMVß-gal plasmid were cotransfected in all experiments as an internal standard. Bars represent the mean luciferase activity ± S.E.M. of at least three independent experiments.

what was observed in Cos7 cells, the pAdMLARE reporter shows a certain level of transcriptional activity by itself (about 7-8 folds compared to the control plasmid pAdMLluc). Such basal activity is instead totally absent in the case of pAdMLR3.

As expected, none of the HOX proteins alone shows significant transcriptional activity on the two reporters (2-3 fold activation on pAdMLARE and no activity at all on pAdMLR3). HOXB1 is able to transactivate pAdMLARE in cooperation with Pbx1 at considerably high levels (about 20 folds compared to the reporter basal level, and 140 folds if calculated over the control plasmid), while the transcriptional activity by HOXA1/Pbx1 and HOXB2/Pbx1 complexes is much weaker (4 folds and 6 folds, respectively, over the reporter basal level). These data differs from the one obtained in Cos7 cells, were the transcriptional activity on the b-1ARE was comparable for all the three A1, B1 and B2 HOX/Pbx1 complexes. On the pAdMLR3 reporter, in the presence of Pbx1, HOXB2 activates transcription around 15 folds while HOXA1 and HOXB1 lead to a transcriptional activation which is more than 100 folds compared to the reporter basal level.

These results show that the transcriptional activity of the HOX/Pbx1 complexes is much stronger in P19 compared to Cos7 cells, where their transactivating potential reaches a maximum of 6-8 folds on both pAdMLARE and pAdMLR3 reporters (compare Fig. 2.1 and Fig. 2.9). This enhancement is restricted to HOX proteins belonging to paralogous group 1 (i.e. HOXB1 and HOXA1), while it is not observed with HOXB2. Moreover, HOXA1 and, even though to a lower extent, HOXB2, cooperatively activate transcription with Pbx1 only through the trimeric R3 site, while the native b-1ARE element seems to be much less permissive for their activity in this cell background.

Cotransfection experiments of either the pAdMLARE or the pAdMLR3 reporter plasmids with HOXB1 and Pbx1 expression vectors performed in HeLa

and NIH3T3 cells (epithelial human carcinoma and immortalized mouse fibroblasts, respectively) led to transactivation levels comparable to those obtained with Cos7 cells, while, if tested in NT2/D1 cells, a cloned human embryonal carcinoma cell line with the characteristic of a neuroectodermal progenitor, the transactivation levels were similar to those observed in P19 (data not shown).

2.7 MUTATIONAL ANALYSIS OF THE b-1ARE ENHANCER.

In order to analyse the specific contribution of each of the b-1ARE conserved repeats in directing transcriptional activity, and at the same time to try to understand the lack of activity of HOXA1/Pbx1 on the b-1ARE element in P19 cells, we performed transient co-transfection experiments in both P19 and Cos7 using reporter constructs containing progressive deletions of the b-1ARE 5' region (Fig. 2.10B). The constructs are schematically represented in Fig. 2.10A. In pAdMLΔ1-52, all the region upstream the three Pbx1 consensus repeats has been removed. In pAdMLΔR1 and pAdMLΔR1+R2 the deletion extends to R1 and to both R1 and R2, respectively. In addition to the three Pbx consensus binding sites, within the b-1ARE sequence there is another stretch of conserved bases which contains a TAAT motif (see Fig. 1.6). This element (ATGCTAAT) is homologous to the recognition sequence of the octamer-binding proteins, another family of transcription factors (Ryan and Rosenfeld, 1997). To test the role of this conserved region in the b-1ARE context, we also generated a reporter construct with a mutated octamer site (pAdMLoctm in Fig. 2.10A) and tested its activity in cotransfection experiments. The results are shown in Fig. 2.10B: in P19, the removal of 5' portions of the b-1ARE progressively decreases the activity of the HOXB1/Pbx1 complex by 30% in

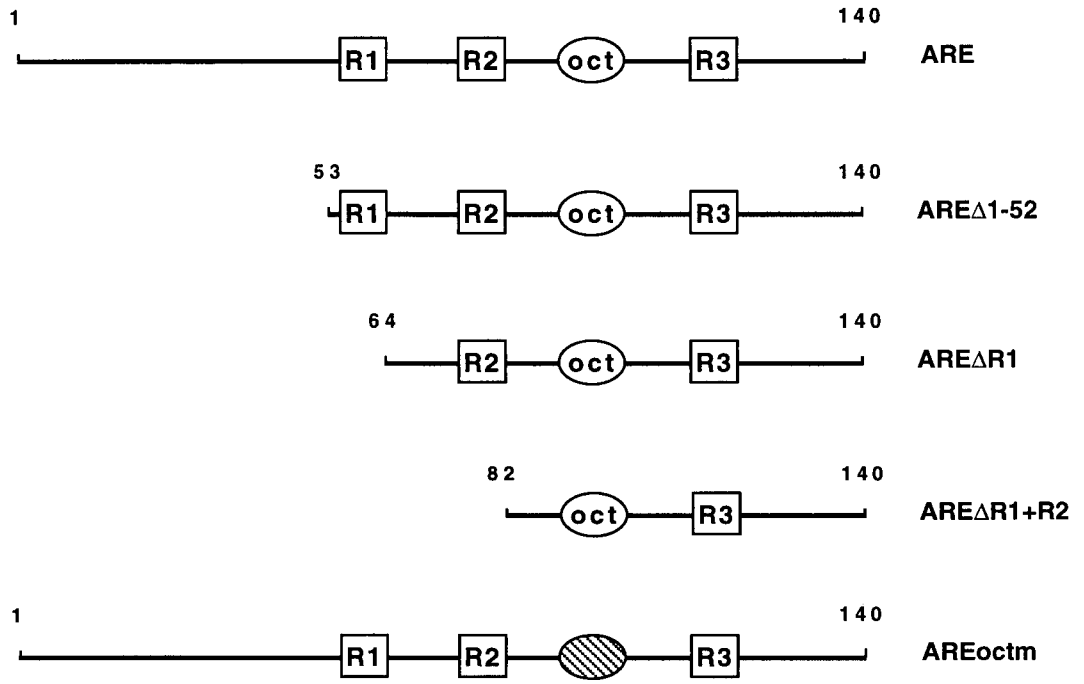


Figure 2.10 A. Schematic representation of the modified b-1ARE versions. R1, R2, R3: Pbx1 consensus repeats. Oct represents the octamer binding site and the striped pattern its relative mutation. All the fragments have been cloned in the pAdMLluc vector.

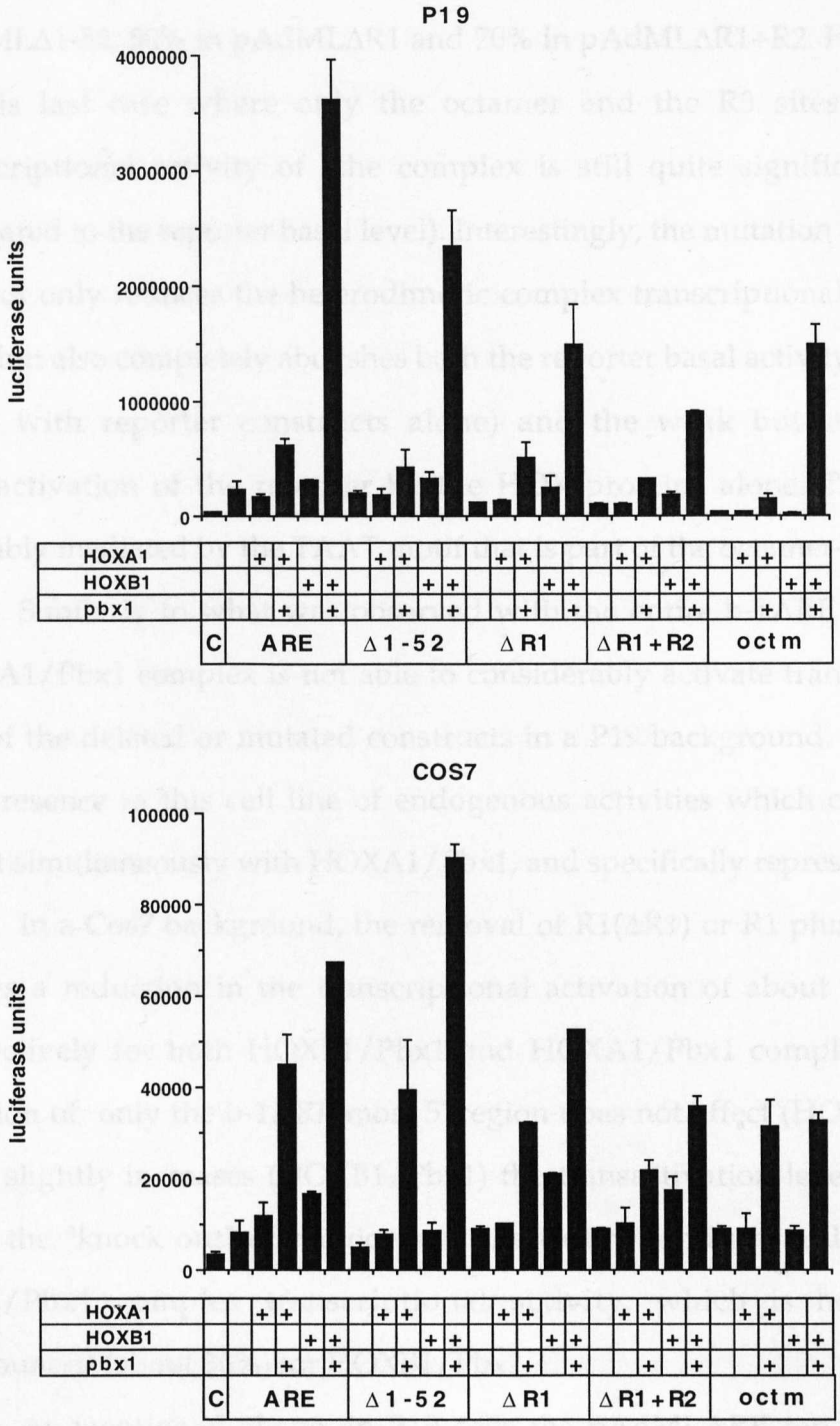


Figure 2.10 B. Luciferase activity from transiently transfected Cos7 and P19 cells. 8 μg of the reporters shown in Fig 2.10 A were cotransfected together with 8 μg of the Pbx1a and 4 μg of HOXA1 and HOXB1 expressors where indicated. 0.2 μg of the pCMVß-gal plasmid were cotransfected in all experiments as an internal standard. Bars represent the mean luciferase activity ± S.E.M. of at least three independent experiments.

pAdMLΔ1-52, 50% in pAdMLΔR1 and 70% in pAdMLΔR1+R2. However, even in this last case where only the octamer and the R3 sites are left, the transcriptional activity of the complex is still quite significant (10 folds compared to the reporter basal level). Interestingly, the mutation in the octamer site not only reduces the heterodimeric complex transcriptional activation by 50%, but also completely abolishes both the reporter basal activity (compare the lanes with reporter constructs alone) and the weak but still detectable transactivation of the reporter by the HOX proteins alone. This activity is probably mediated by the TAAT motif that is part of the octamer-like site.

Similarly to what was observed with the entire b-1ARE enhancer, the HOXA1/Pbx1 complex is not able to considerably activate transcription with any of the deleted or mutated constructs in a P19 background. This rules out the presence in this cell line of endogenous activities which could bind the target simultaneously with HOXA1/Pbx1, and specifically repress it.

In a Cos7 background, the removal of R1(ΔR1) or R1 plus R2 (ΔR1+R2) causes a reduction in the transcriptional activation of about 20% and 40% respectively for both HOXB1/Pbx1 and HOXA1/Pbx1 complexes, whereas deletion of only the b-1ARE most 5' region does not affect (HOXA1/Pbx1) or even slightly increases (HOXB1/Pbx1) the transactivation level. Also in this case, the "knock out" of the octamer site (octm) produces a decrease in the HOX/Pbx1 complex transcriptional activity, which is however more pronounced (about 30%) for HOXB1/Pbx1.

As mentioned above, in P19 cells the pAdMLARE reporter construct displays a certain level of basal activity on its own (around 8 fold) compared to the control plasmid pAdMLluc. Such activity does not depend on the presence of endogenous HOX/Pbx complexes, since it is not detected with the reporter containing a trimer of the R3 repeat (pAdMLR3, Fig. 2.9), but is probably due to other factors binding on the octamer site, and it is completely abolished with

the reporter in which the octamer has been disrupted (pAdMLoctm, Fig. 2.10B). The pAdMLARE reporter has no basal activity in Cos7 cells, favouring the hypothesis that the octamer site in Cos7 and P19 cells could be bound, at least in part, by different proteins. Nevertheless, the mutational analysis indicates the presence in both cell lines of an octamer-dependent activity which sinergistically cooperates with Hox/Pbx heterodimers in transactivating the b-1ARE reporter. This suggests a role for this conserved octamer-like site, in addition to the three Pbx repeats, in regulating the function of the b-1ARE enhancer.

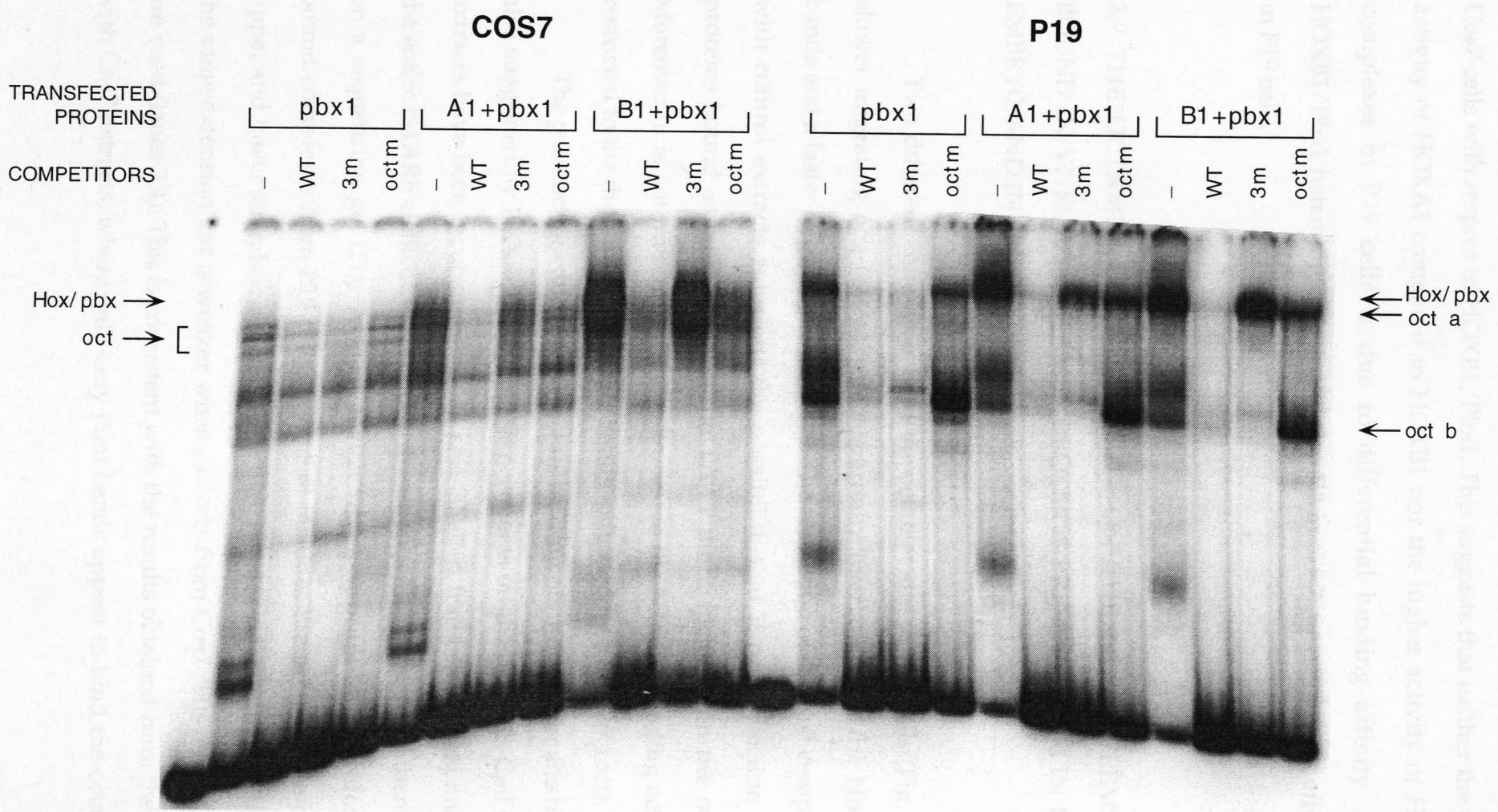
The higher transcriptional activity of the HOXB1/Pbx1 complex in P19 compared to the one obtained in Cos7 cells is still quite evident on the pAdMLoctm reporter (50 versus 4 folds over the reporter basal level in P19 and Cos7, respectively) (Fig. 2.10B). In this case, and the same applies to pAdMLR3, since there is no basal reporter activity when compared to the control plasmid pAdMLluc (C), and no transactivation by the HOX protein alone, the transcriptional activity observed when Pbx1 and either HOXA1 or HOXB1 are cotransfected must derive entirely from the HOX/Pbx1 complex working on the three Pbx repeats. This means that the stronger transactivation observed in P19 does not depend on the contribution by additional transcription factors binding to the octamer site, but to a different effector function of the complex in this particular cell background. That this effect is not a general property of all transcription factors in the EC cell system is demonstrated by the fact that the transcriptional activity of a VP16-Gal4 chimeric protein (containing the activation domain of the Herpes Simplex Virus activator protein VP16) tested on a Gal4 reporter construct is comparable in Cos7 and P19 cells (data not shown).

2.8 THE DIFFERENT TRANSCRIPTIONAL ACTIVITY OF HOXA1/PBX1 AND HOXB1/PBX1 COMPLEXES IN P19 AND COS7 CELLS DOES NOT DEPEND ON DIFFERENTIAL BINDING EFFICIENCY.

To test whether a different cell context could affect the binding of HOX/Pbx complexes to the b-1ARE target, we performed electrophoretic mobility shift assays (EMSA) with extracts from transfected Cos7 and P19 cells, using the ΔR1+R2 deletion mutant of the ARE element as a probe. This fragment (60 base pairs), contains the R3 site together with the adjacent octamer consensus sequence (AREΔR1+R2 in Fig. 2.10A).

Total extracts prepared from cells which had been previously co-transfected with either Pbx1 and HOXA1 or Pbx1 and HOXB1 expression plasmids, were incubated with the probe in the absence ("-" in Fig. 2.11) or in the presence of a 200-fold excess of the indicated unlabelled competitor fragment. The mix was then loaded on a polyacrylamide gel to separate retarded complexes. Extracts from untransfected cells give the same binding pattern (not shown) of extracts from cells transfected with the Pbx1 expression vector alone, which were used as a control. The results are shown in Fig. 2.11: the slowest mobility complexes present in co-transfected cells represent Hox/Pbx1 heterodimers bound on the R3 site, as confirmed by competitions. The binding efficiency of HOXA1/Pbx1 and HOXB1/Pbx1 complexes is comparable in P19 cells, while the HOXA1/Pbx1 seems to bind very poorly in

Figure 2.11. EMSA analysis of total extracts from transfected Cos7 and P19 cells on the AREΔR1+R2 fragment (described in Fig. 2.10A). Cells were transfected with 10 μg of the indicated expression construct and the extracts (5-10 μg per lane) were normalized for both protein concentration and transfection efficiency. A 200 fold excess of unlabelled competitor fragment was added to the binding mix where indicated: WT= AREΔR1+R2 wild type; 3m= AREΔR1+R2 with a mutation in the R3 site; octm= AREΔR1+R2 with a mutation in the octamer site. All competitor fragments were produced by PCR amplification of the b-1ARE element with appropriate primer oligos. Bands corresponding to HOX/Pbx1 and octamer binding complexes (oct) are indicated by arrows.


COS7
P19
TRANSFECTED PROTEINS
pbx1
A1+pbx1
B1+pbx1
COMPETITORS
–
WT
3m
oct m
Hox/ pbx
oct
Hox/ pbx
oct a
oct b

Cos7 cells with respect to HOXB1/Pbx1. This suggests that neither the reduced activity of HOXA1 compared to HOXB1 nor the higher activity of Hox/Pbx complexes in P19 cells is due to differential binding affinity. In fact, HOXB1/Pbx1 heterodimer formation on R3 in Cos7 seems to be as efficient as in P19 cells.

2.9 THE OCTAMER CONSENSUS SITE ON THE b-1ARE ELEMENT IS BOUND IN VITRO BY OCTAMER PROTEINS PRESENT BOTH IN MOUSE EMBRYO AND IN EMBRYONAL CARCINOMA CELL LINES.

Two different complexes bind the octamer site in P19 cells (Fig. 2.11): a slower migrating one (oct a), which partially overlaps with the Hox/Pbx1 bands and a faster migrating one (oct b). The formation of this last complex also with control extracts rules out the possibility that it could be due to Hox proteins bound as monomers on the TAAT motif included in the octamer. Moreover, it is not present in Cos7 cells, where the octamer binding activity is restricted to the three slow migrating bands indicated by a bracket (oct).

The presence of endogenous factors binding on the octamer site has been also suggested by DNase I footprints experiments on the b-1ARE. Cell nuclear extracts have been incubated with an end-labelled fragment corresponding to the entire b-1ARE element, subjected to digestion with DNase I and then loaded on a sequencing gel (Fig. 2.12). A region which appears to be protected by bound complexes from P19 cells (lanes 5,6) includes the octamer site on both upper and lower strands (right and left panel, respectively). The protection has the same extension but is weaker when extracts from Cos7 cells instead of P19 are used (lanes 3,4). This is consistent with the results obtained from the EMSA with Cos7 extracts, where only very faint bands appear to bind the octamer in

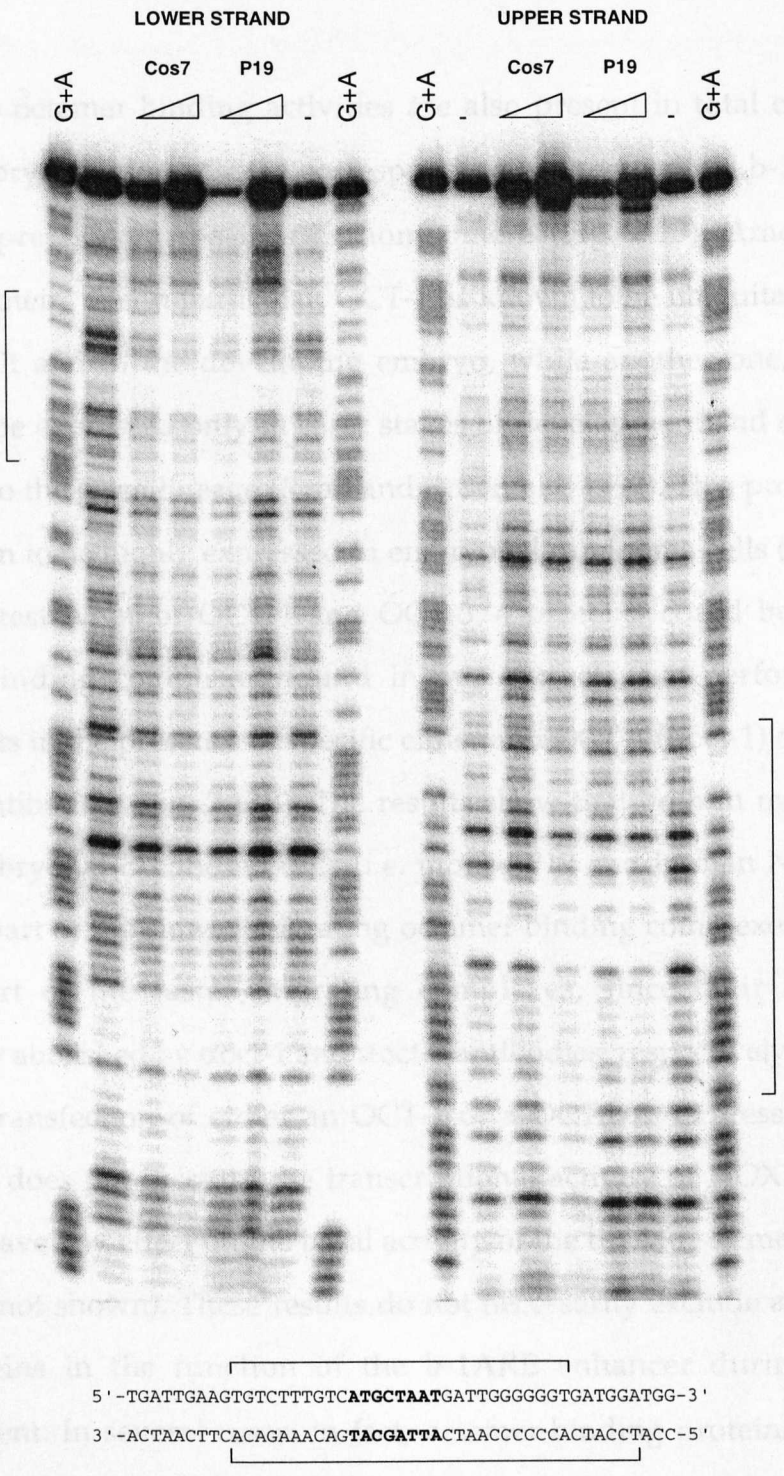


Figure 2.12. DNase I footprinting analysis of the 140-bp b-1ARE enhancer element. End labelled DNA fragments on either upper or lower strand were incubated with 10 or 20 µg of nuclear extract from Cos7 or P19 cells before digestion with DNase I. G+A= Maxam and Gilbert sequence reaction of the probe on G and A residues. The sequence of the protected region (brackets) is shown below as a shaded box. The octamer site is in bold.

contrast with the very abundant octamer binding activity present in P19 cells (Fig. 2.11).

Two octamer binding activities are also present in total extracts from mouse embryo 9.5 dpc, a stage corresponding to the maximal b-1ARE-driven *Hoxb-1* expression in the fourth rhombomere (Fig. 2.13). Among octamer binding protein, one in particular, OCT-1, is known to be ubiquitary expressed in the adult and in the developing embryo, while another one, OCT3/4, is known to be expressed only in early stages of development and subsequently restricted to the germ lineage (Ryan and Rosenfeld, 1997). This protein has also been shown to be highly expressed in embryonal carcinoma cells (Viganò *et al.*, 1996). To test whether OCT-1 and OCT-3/4 proteins could be part of the octamer binding complexes found in our extracts, we performed EMSA experiments in the presence of specific either anti OCT-1 (αoct-1) or anti OCT-3 (αoct-3) antibodies (Fig. 2.13B). The results show that both in mouse embryo and in embryonal carcinoma cells (i.e. mouse P19 and human NT2) extracts, OCT-1 is part of the slower migrating octamer binding complexes while OCT-3/4 is part of the faster migrating complexes, since their formation is completely abolished by αoct-1 and αoct-3 antibodies, respectively.

Cotransfection of either an OCT-1 or a OCT3/4 expression vector in Cos7 cells does not increase the transcriptional activity of HOXB1/Pbx1 and does not have any effect on the basal activity of the b-1ARE element in this cell line (data not shown). These results do not necessarily exclude a role of these two proteins in the function of the b-1ARE enhancer during hindbrain development. In several cases, in fact, octamer-binding proteins activity has been shown to be regulated by cell context-dependent events, such as protein phosphorylation, which influences their DNA binding activity, and the interaction with tissue-specific cofactors, which modulate their transcriptional activation functions (reviewed in Ryan and Rosenfeld, 1997). The requirement

for a specific cell background could then, at least in part, account for the inability of transfected OCT-1 and OCT3/4 to activate transcription through the b-1ARE enhancer in Cos7 cells.

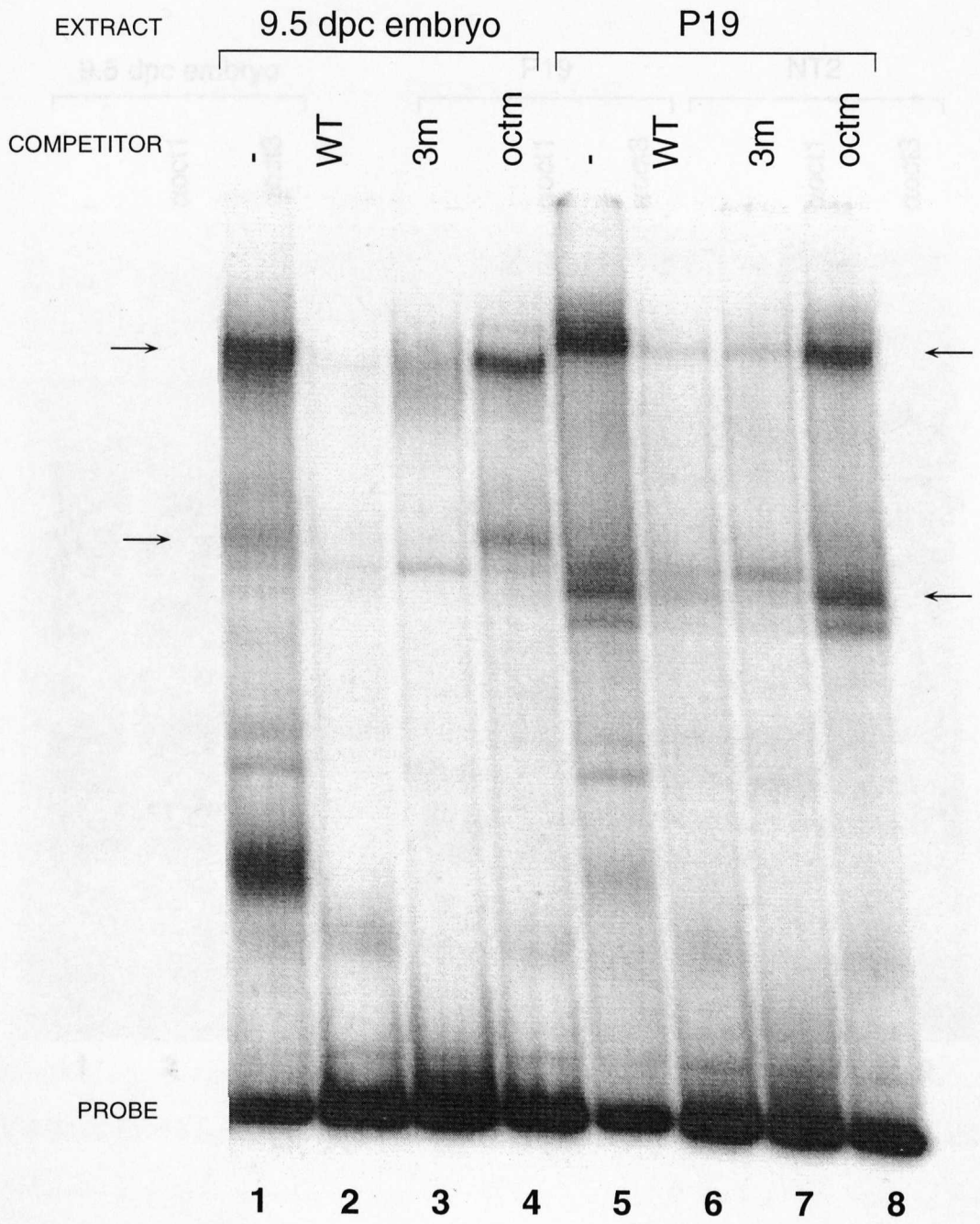


Figure 2.13 A. EMSA analysis of total extracts (about 10 µg per lane) from 9.5 dpc mouse embryos and embryonal carcinoma cells on the AREΔR1+R2 fragment. Complex formation was challenged by adding a 200 fold excess of the indicated unlabelled competitor fragment. Arrows indicate octamer bound proteins.

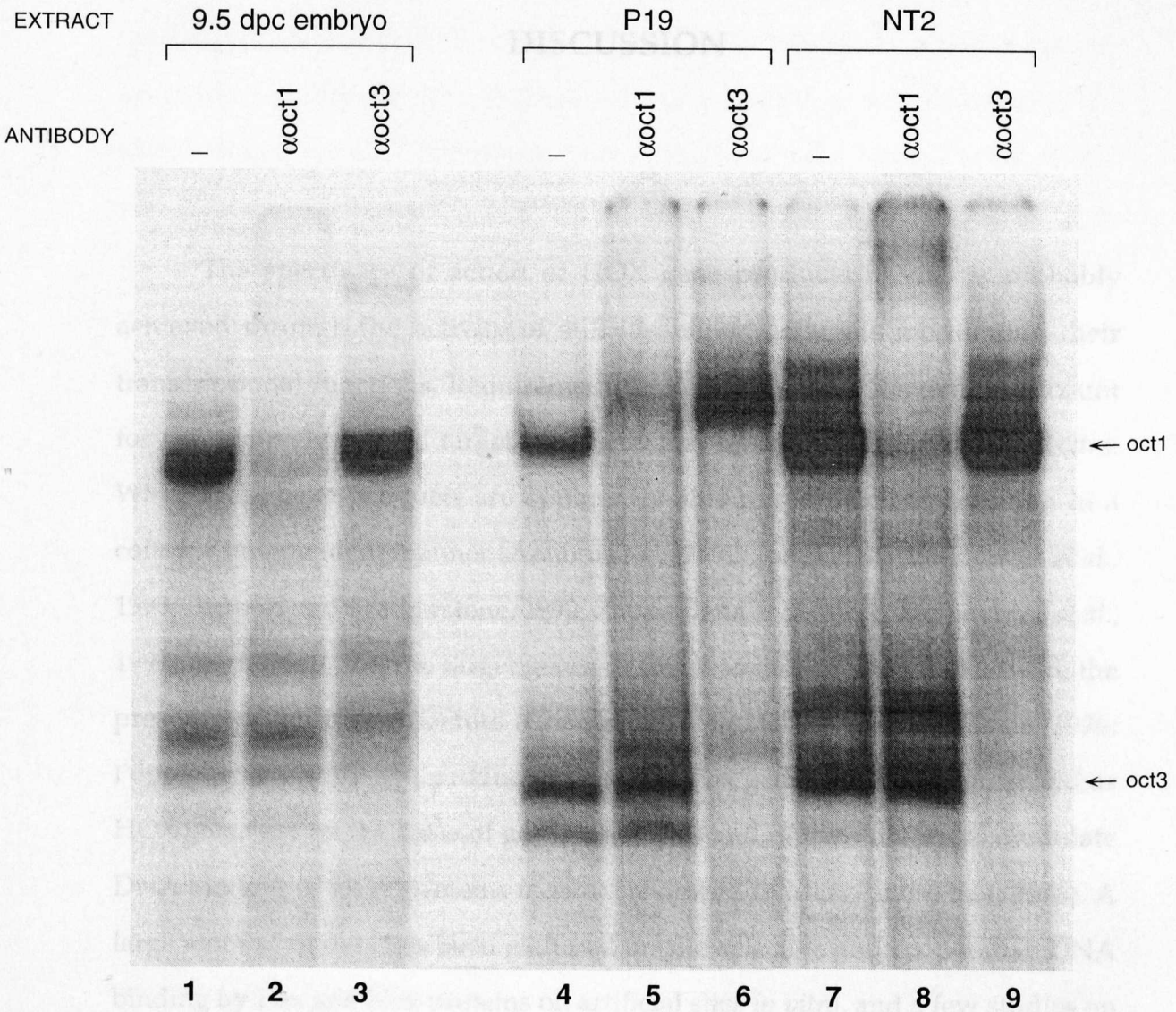


Figure 2.13 B. EMSA analysis of total extracts (about 10 µg per lane) from 9.5 dpc mouse embryos and embryonal carcinoma cells on the AREΔR1+R2 fragment. Complex formation was challenged by adding 2 µl of the indicated antibody. Arrows indicate octamer bound proteins.

Chapter 3

DISCUSSION

The specificity of action of HOX gene products *in vivo* is probably achieved through the activity of still ill-defined cofactors modulating their transcriptional functions. Requirement for such factors *in vivo* would account for the relatively relaxed target specificity displayed by HOX proteins *in vitro*. While HOX gene products are apparently able to regulate transcription in a cofactor-independent manner (Arcioni *et al.*, 1992; Jones *et al.*, 1992; Jones *et al.*, 1993; Pöpperl and Featherstone, 1992; Zappavigna *et al.*, 1991, Zappavigna *et al.*, 1994), regulation of some target genes *in vivo* was indeed shown to require the presence of additional factors (Chan *et al.*, 1994; Gross and McGinnis, 1996; Pöpperl *et al.*, 1995). The products of the *exd*/Pbx genes have been proposed as HOX cofactors on the basis of genetic analysis and of their ability to modulate DNA binding of HOX proteins *in vitro* (reviewed in Mann and Chan, 1996). A large amount of data has been gathered on site-selective and cooperative DNA binding by Pbx and Hox proteins on artificial sites *in vitro*, and a few studies on the E2A-Pbx oncogene fusion provided evidence for functional cooperativity between the two families of proteins (Chang *et al.*, 1996; Lu and Kamps, 1996a; Lu *et al.*, 1995; Peers *et al.*, 1995; Phelan *et al.*, 1995). Nevertheless, the complex formed by wild-type Pbx and Hox proteins had never been functionally characterized in terms of its ability to cooperatively activate transcription from known target sequence.

The results described in this thesis show that Pbx1 can modulate HOX protein function at the level of transcription, by acting as a cofactor in the activation of a target element by a restricted class of human HOX proteins. We used as a model a genetically characterized, Pbx-dependent Hox target, the 148-bp autoregulatory element (b-1ARE) driving *Hoxb-1* gene expression in the fourth rhombomere of the mouse developing hindbrain (Pöpperl *et al.*, 1995). We show that the activity of a reporter construct containing the b-1ARE can be stimulated by the HOXB1 protein in transfected cells only upon coexpression of Pbx1. The b-1ARE contains three repeated sequences closely related to an *in vitro*-selected Pbx1 consensus binding site (Pöpperl *et al.*, 1995; Van Dijk *et al.*, 1993), one of which (R3) was previously shown to bind *in vitro* a *Hoxb-1/exd* complex, and to be crucial for the *in vivo* activity of the b-1ARE (Pöpperl *et al.*, 1995). Consistent with these findings, transactivation of the b-1ARE by the HOXB1/Pbx1 complex is significantly reduced by a mutation within the R3 site, while a multimerized R3 site is sufficient to mediate cooperative activation in a promoter- and cell context-independent fashion. Thus, HOXB1 requires Pbx1 as a transcriptional cofactor to regulate the activity of the b-1ARE through the R3 site.

Coexpression of Pbx1 with a representative variety of HOX proteins showed that the b-1ARE, or the multimerized R3 site, is transactivated only by members of the paralogy groups 1 and 2, namely HOXA1, HOXB1 and HOXB2. The fact that the b-1ARE is activated by other proteins besides HOXB1, may reflect the existence of cross-regulation of this element *in vivo* by a subset of Hox genes. It is noteworthy, in this respect, that *Hoxa-1* and *Hoxb-2* are both expressed in rhombomere 4 during development (reviewed in Keynes and Krumlauf, 1994), and that ectopic expression of *Hoxa-1* was shown to cause activation of a *Hoxb-1-lacZ* reporter construct in the hindbrain of transgenic mice (Zhang *et al.*, 1994). This activation, in fact, does not simply reflect a

functional redundancy among paralogous group-1 Hox proteins. Although *Hoxb-1* is expressed normally in rhombomere 4 in *Hoxa-1* null mouse mutants (Carpenter *et al.*, 1993; Mark *et al.*, 1993), it has been recently shown that *Hoxa-1* and *Hoxb-1* work sinergistically on the b-1ARE enhancer in initiating the r4-restricted expression of *Hoxb-1* (Studer *et al.*, 1998). Nevertheless, expression of a b-1ARE-*lacZ* reporter in *Drosophila* embryos is not entirely dependent on the function of the *Drosophila* gene *labial* (Pöpperl *et al.*, 1995), suggesting that other Hox genes could indeed participate to the regulation the b-1ARE in a cofactor-dependent manner.

The selectivity of the *Hoxb-1* ARE, and particularly of the R3 site, appears to depend on its ability to allow the assembly of a Hox/Pbx1/DNA ternary complex only with group 1 and 2 Hox proteins, as indicated by DNA binding analysis *in vitro*. These results underline the role of Pbx1 as a cofactor able to confer DNA recognition specificity to Hox proteins: in the presence of Pbx, only a small subset of Hox proteins is able to recognize a particular target site. These results parallel those obtained *in vivo*, since only the Hox/Pbx complexes which recognise the R3 site are able to transactivate the reporter in transfected cells. The experiments reported by Chan, which show that changing the two central basepairs in the R3 site is sufficient to switch the *Hox/exd* dependent expression pattern of a reporter construct in *Drosophila* from *labial* to *Deformed* (Chan et *al.*,1997), are totally complementary to our results and, together with them, provide *in vivo* evidence that different Hox/Pbx complexes may regulate transcription through different target sites.

The cofactor-mediated, differential site recognition of HOXB1 appears to be specified by the extended N-terminal region of the Hox homeodomain. In fact, the ability to complex with Pbx1 on the R3 site, and to activate transcription through the b-1ARE, can be transferred to the HOXB3 protein by

swapping a region containing only the homeodomain N-terminal arm. This result is in agreement with previous studies showing that the flexible homeodomain N-terminal region is crucial for determining the specificity of action of Hox proteins *in vivo* (Chan *et al.*, 1994; Furukubo-Tokunaga *et al.*, 1993; Lin and McGinnis, 1992; Zappavigna *et al.*, 1994; Zeng *et al.*, 1993), and with previous models derived by *in vitro* studies on PCR-selected targets suggesting that the "YPWM"/homeodomain N-terminal region is an important determinant in the site-selectivity of different Hox/Pbx complexes *in vitro* (Chan *et al.*, 1996; Chang *et al.*, 1996; Lu and Kamps, 1996b). In the context of the HOXB1/Pbx1 complex, the "YPWM" motif therefore provides only a neutral, protein-protein interaction function, as previously suggested by yeast two-hybrid experiments (Johnson *et al.*, 1995), without necessarily contributing to the DNA binding specificity. Indeed, mutations in the domains involved in protein-protein contacts, either in HOXB1 (deletion or mutation of the critical tryptophan in the FDWM region) or in Pbx1 (deletion of the C-terminus downstream from the homeodomain), (see Chang *et al.*, 1996; Lu and Kamps, 1996b), disrupt the formation of the ternary complex *in vitro* and abolish transcriptional activation altogether. Furthermore, we show that transcriptional activation by the HOXB1/Pbx1 complex requires intact DNA binding functions by both HOXB1 and Pbx1 proteins. Thus, unlike the MATa1/MATα2 homeodomain protein complex, where a MATα2 mutation impairing DNA binding is still able to form a functional complex with MATa1 on DNA *in vivo* (Vershon *et al.*, 1995), in the case of the HOXB1/Pbx1 complex tethering of either protein onto DNA through protein-protein interactions is not sufficient for activity.

The transcriptional activity of the HOXB1/Pbx1 complex on the b-1ARE was exploited to characterize functional domains in both proteins. The main

transcriptional activation domain within the complex was mapped to the N-terminal region of the HOXB1 protein. In fact, deletion of the 155 N-terminal amino acids of the HOXB1 protein considerably reduced the transcriptional activation strength without affecting cooperative DNA binding. The N-terminus of the HOXB1 protein activates transcription also when fused to an heterologous DNA binding domain. However, deletion of the HOXB1 N-terminal domain did not completely abolish transcriptional activity, leading us to speculate that Pbx1 could also contribute to the activation strength of the complex. A transcriptional activation domain was indeed mapped within the 83 C-terminal amino acids of the Pbx1 protein. This region is absent in the Pbx1b alternative splicing variant, which reproducibly displayed lower levels of activation when compared to the full-length protein (Pbx1a) in combination with HOXB1. The C-terminal region of Pbx1a is rich in Ser/Thr residues, and is able to activate transcription when fused to a heterologous DNA-binding domain. Interestingly, a Ser/Thr rich region located within the C-terminus of the LFB1 (HNF1) homeodomain protein has been previously reported to be implicated in transcriptional activation, and to account for 50% of the LFB1 activity (Nicosia *et al.*, 1990).

The C-terminal region of Pbx1a is apparently not functional as an activation domain in the uncomplexed Pbx1a protein, which is unable to activate transcription as a monomer even through a PCR-derived optimal consensus binding site (Mann and Chan, 1996; Van Dijk *et al.*, 1993). Deletion of the N-terminal 230 amino acids, containing the two conserved PBC-A and PBC-B regions (Bürglin and Ruvkun, 1992), uncovers the function of the Pbx1 C-terminal activation domain, since the Pbx1Δ1-230 mutant is able to significantly stimulate the b-1ARE R3 reporter activity even in the absence of HOXB1. A deletion of the N-terminal 232 amino acids of Pbx1 has been previously reported to bind to a Pbx consensus sequence with higher affinity with respect

to its wild-type counterpart in the absence of Hox proteins (Lu and Kamps, 1996b), suggesting that the activity of the Δ1-230 mutant on the R3 reporter could be due to an increased binding affinity for the R3 site. Transcriptional activation by the Pbx1aΔ1-230 mutant is sustained by the C-terminal activation domain, since the same N-terminal deletion in the shorter Pbx1b variant, Pbx1bΔ1-230, does not significantly activate transcription. The Pbx1b alternatively spliced variant may therefore represent a less active form of Pbx1, capable of forming a DNA-bound complex with HOX proteins like the longer Pbx1a variant, but leading to lower levels of transcriptional activation. Pbx1b could antagonize and/or substitute for the longer and more active Pbx1a isoform in different tissues or body regions, allowing a fine tuning of the activity of Hox/Pbx complexes on their targets.

The constitutive expression of *Hoxa-1* and *Hoxb-1* in transgenic mice was shown to induce the expression of an *Hoxb-1*-lacZ reporter construct in an ectopic region, anterior to the normal r4 expression domain, which is localized and restricted to r2 (Zhang *et al.*, 1994; Pöpperl *et al.*, 1995). Since Pbx genes are widely expressed in embryonic and adult tissues, this spatially restricted induction by constitutively expressed Hox proteins implies that additional factors are either necessary for, or interfere with, the function of the Hox/Pbx complexes. We found that the transcriptional activity of the Hox/Pbx1 complexes is much higher (more than 10 times) in embryonal carcinoma (EC) cells than in other cell lines. Murine and human EC cells can be induced to differentiate *in vitro* by retinoic acid (RA) into neurons and astrocytes, while in the presence of dimethyl sulphoxide (DMSO) they differentiate into a spectrum of cells which include cardiac and skeletal muscle. Such differentiated cells types resemble cellular populations present in the extraembryonic and embryonic portions of the developing mouse and both *in vivo* and *in vitro*

studies suggest that the developmental sequence of events which occurs in the differentiating population of the EC cells is very similar to the analogous events occurring in the embryo (Rudnicki and McBurney, 1987; Simeone *et al.*, 1990). For this reasons, EC cells could represent a more suitable or "competent" background for the functioning of Hox/Pbx1 complexes and the higher transcriptional activity observed in these cells, - although differences due to post translational modifications of the two proteins in different contexts cannot be ruled out -, could reflect the presence of additional, specific factors which cooperate with Hox/Pbx to activate transcription. The existence of cell-type or tissue restricted cofactors that act in concert with HoxB1/Pbx1 could as well help to explain the spatially limited induction of the autoregulatory response observed *in vivo*.

Deletion of R1 or of R1 and R2 repeats from the b-1ARE element, causes a progressive decrease of transcriptional activation by Hox/Pbx complexes both in Cos7 and in P19 cells. These results are in agreement with the in vivo data showing that, although the R3 site is the most effective one, all three repeats contribute to enhancer activity in the mouse (Pöpperl *et al.*, 1995).

In addition to the three Pbx repeats, we showed that also an octamer-like element adjacent to the R3 site contributes to the transcriptional activity directed by the b-1ARE enhancer, most likely through the binding of factors which act in synergy with Hox/Pbx heterodimers. The octamer cis-acting transcriptional regulatory motif (consensus sequence ATGCAAAT) is found in enhancers and promoters of many genes, expressed either ubiquitously or in a tissue specific fashion. This motif regulates gene expression by the binding of transcription factors belonging to the POU family (reviewed in Ryan and Rosenfeld, 1997). These factors are characterized by the presence of the POU domain, a bipartite DNA binding structure containing an N-terminal POU

specific region (POUs) and a POU homeodomain (POUhd) connected by a linker region. In addition to their DNA binding functions, the POU domain can participate in protein-protein interactions with both POU proteins and other transcriptional regulators. Homo- and heterodimerization, mediated by the POU domain, has been demonstrated for several POU proteins binding to multiple adjacent sites (Herr and Cleary, 1995; Ryan and Rosenfeld, 1997). By EMSA analysis, we show that two POU proteins, namely Oct1 and Oct3/4, bind to the octamer motif of the b-1ARE enhancer in EC cells. Of course, we do not have at the moment any evidence indicating an interaction with Hox/Pbx complexes of these two factors or their direct involvement in regulating transcription by the b-1ARE. However, some data reported on a zebrafish POU protein, Pou-2, whose expression is restricted to r2 and r4 during hindbrain development (Hauptmann *et al.*, 1995), are consistent with the hypothesis of a possible involvement of this class of proteins in the establishment or maintaining of r2 and r4 segments, maybe by participating in the control of Hox genes expression.

In addition to the enhanced transcriptional activity, the cotransfection experiment performed in P19 cells uncovered an additional functional restriction for different Hox/Pbx1 complexes which was not evident in Cos7 cells. HOXA1, HOXB1 and HOXB2 are in fact all able to form heterodimers with Pbx1 on the R3 site and in Cos7 cells they cooperatively activate transcription with Pbx1 at comparable levels both through the entire b-1ARE element and the trimeric R3 site. However, when tested in P19, there were substantial differences in their transactivating potential. In P19, the HOXB1/Pbx1 complex is able to activate transcription at high levels from both pAdMLARE and pAdMLR3 reporter constructs, while the HOXB2/Pbx1 complex exhibits a very weak transcriptional activity on both reporters. This is

consistent with *in vivo* data showing that ectopic expression of *Hoxb-2* in transgenic mice is unable to induce the expression of a reporter gene through an Hoxb-1/Pbx target site (Maconochie *et al.*, 1997). It is therefore possible that the transcriptional function of HOXB2/Pbx1 in the embryo, as well as in EC cells, requires a different target configuration, which is not reproduced in the b-1ARE element.

The results obtained with HOXA1 are quite intriguing: in P19 cells, HOXA1/Pbx1 is able to activate transcription as efficiently as HOXB1/Pbx1 from the trimeric R3 site, while it is almost inactive on the b-1ARE target. When tested on the pAdMLΔR1+R2 construct, which contains only the octamer-like and the R3 sites, HOXB1/Pbx1 is still able to strongly activate transcription, while HOXA1/Pbx1 is not. Nevertheless, as shown by EMSA experiments where the ΔR1+R2 fragment has been used as a probe, in extracts from transfected P19 cells, HOXA1/Pbx1 and HOXB1/Pbx1 heterodimers bind the R3 site with comparable affinity. A reduced or different binding affinity of the HOXA1/Pbx1 complex for one of the three repeats, compared to HOXB1/Pbx1, is therefore not sufficient to explain their different behaviour. A possible inhibitory effect of the octamer-like site has also been ruled out, since its mutation does not rescue HOXA1/Pbx1 transcriptional activity. A different activity of HOXA1 and HOXB1 on b-1ARE could then reside in a different ability of the two proteins to interact with, and hence to recruit, other cofactors to the transcriptional machinery. The need to contact other cofactors for reaching high transcriptional activation levels could be overcome in the case of the multimerized pAdMLR3 reporter, where the possibility to form multiple Hox/Pbx complexes working in synergy on three adjacent R3 sites could compensate for a weaker transactivation capacity.

As outlined above, the transcriptional activating functions of *Hoxa-1* and *Hoxb-1* on the b-1ARE enhancer are not completely redundant, since both genes

are required for the correct initiation of *Hoxb-1* expression in r4 (Studer *et al*, 1998). The different behaviour of the products of these two genes on a specific target (the b-1ARE) in different cellular contexts, could then indeed reflect their functional diversity.

Since the homeodomains are almost identical, the specificity of function of HOXA1 and HOXB1, and of paralogous genes in general, must depend on regions outside the homeodomain, where the proteins completely diverge. The utility of new genes that arose by duplication of Hox clusters is to provide the developmental plasticity to evolve more complex body plans, and may have depended on their ability to acquire new regulatory specificities and hence new functions. This could be achieved, at least in part, by retaining similar DNA binding specificities by conserving the homeodomain, while allowing other regions of the proteins to diverge to the point where they could interact with different cofactors to achieve distinct developmental outcomes.

Chapter 4

FUTURE PROSPECTS

Recently other homeodomain containing proteins have been identified that are able to form complexes with Exd and Pbx. These proteins include the *Drosophila* Homothorax (HTH), its vertebrate ortholog Meis1 and the human Prep1, and are collectively known as the Meis-related proteins. Like PBC proteins, they belong to the TALE (three aminoacids loop extension) superfamily of homeodomain containing proteins. They share extensive homology in their homeodomain and are characterized by the presence of two highly conserved short domains, namely HR1 and HR2, located in their N-terminal region (reviewed in Mann and Affolter, 1998).

Genetic studies in *Drosophila* provided evidence that HTH functionally interacts with Exd, and that HTH and Exd together act as Hox cofactors during development: HTH has been shown to act as a master antennal determining gene during development (Casares and Mann, 1998), to suppress eye development (Kurant *et al.*, 1998), and to cooperate with Exd in patterning the embryonic peripheral nervous system (Pai *et al.*, 1998). HTH is required for the nuclear localization of Exd: in the absence of HTH, Exd is in the cytoplasm and presumably not functional. HTH and Meis are both able to induce Exd's nuclear translocation and both interact with Exd in the absence of DNA. Therefore, HTH could regulate Exd availability in the nucleus and hence the possibility to interact with Hox proteins. This level of control appears to be reciprocal, since

Hox genes control Exd levels as well, probably regulating *hth* transcription (reviewed in Mann and Affolter, 1998).

Meis1-related proteins are found complexed with Pbx proteins in a variety of cell types. The Prep1-Pbx1 complex has been identified and purified as UEF3 (Urokinase Enhancer Factor 3), a factor involved in the regulation of the Urokinase enhancer and of several other AP1 controlled promoters (Berthelsen *et al.* 1996, De Cesare *et al.*, 1997). Prep1 and Pbx most likely exist as stable complexes in the nucleus, as stable Prep1-Pbx complexes can be isolated from nuclear extracts . Unlike the Pbx/Hox complex, Prep1 and Pbx dimerize efficiently in solution independently of the presence of a DNA target site. Heterodimerization results in a strong DNA binding affinity for the TGACAG Meis target site but also for Pbx-Hox target sequences like the R3 element (Berthelsen *et al.*, 1998).

The interaction between Prep1 and Pbx1 requires sequences in the N-terminus of both proteins, including the PBCA region of Pbx1 and the HR1 and HR2 regions of Prep1, but is independent of the integrity of the homeodomains (Berthelsen *et al.*, 1998). In addition, Prep1 does not contain any YPWM motif, meaning that Pbx forms heterodimers with Prep1 and Hox proteins through totally different interaction surfaces. The use of different interaction surfaces could allow Pbx to interact with Prep1 and Hox proteins simultaneously. Indeed, the formation of a ternary Prep1-Pbx-HoxB1 complex has been observed in vitro on the R3 site. Moreover, although the Prep1-Pbx complex is not able to activate transcription by itself, the addition of Prep1 in cotransfection experiments enhances the transcriptional activity of HoxB1/Pbx1 on the trimeric R3 site. This indicates that the formation of transcriptionally active Hox/Pbx complexes is not antagonized by the presence of inactive Prep1/Pbx heterodimers and suggests a direct functional interaction between HoxB1, Pbx1 and Prep1. Consistent with this hypothesis, Prep1-Pbx complexes

are present in 9.5 dpc mouse embryo, a time when Pbx1 is also interacting with HoxB1 in the developing hindbrain.

Prep1, and Meis related proteins in general, could then be an additional component in the transcriptional regulation by Hox proteins. Moreover, Meis1 is able to form heterodimers with posterior Hox proteins (paralogous group 9-13). Interestingly, the expression of Meis1 and either Hoxa-7 or Hoxa-9 are coactivated by retroviral integration in BXH2 murine myeloid leukaemias (Nakamura *et al.*, 1996).

Considering these new findings, we can represent the ARE enhancer in a possible active configuration which includes octamer binding proteins on the octamer site, and HoxB1/Pbx1/Prep1 on the Pbx repeats as a trimeric complex. The model is shown in Fig. 3.1: HoxB1 stability and activity are influenced by other DNA-binding proteins working on its own as well on adjacent sites. Therefore, although Hox/Pbx binding sites are key elements in Hox-activated enhancers, it is possible that they do not normally function by themselves but need to cooperate with other factors. Other evidences suggesting that this could be effectively the case come from the study of lab550, a Dpp-responsive enhancer of the *Drosophila* HoxB1 ortholog labial. This enhancer requires the binding of labial/Exd heterodimers to an R3-like site to be activated by Dpp (decapentaplegic), a signaling molecule belonging to the superfamily of TGF-β related proteins, suggesting that lab550 integrates inputs from both labial/Exd and Dpp and that these two inputs synergize to activate transcription (Grieder *et al.*, 1997).

In this picture, Hox complexes appear to act in the context of an entire enhancer that integrates a variety of regulatory information from both Hox and signaling pathways. It is likely that Hox complexes synergize with other DNA binding complexes to recruit additional cofactors or the basal transcriptional machinery to ultimately control transcription. Important goals for the future

include dissecting these protein-protein and protein-DNA interactions, and characterizing their relevance to Hox specificity and function.

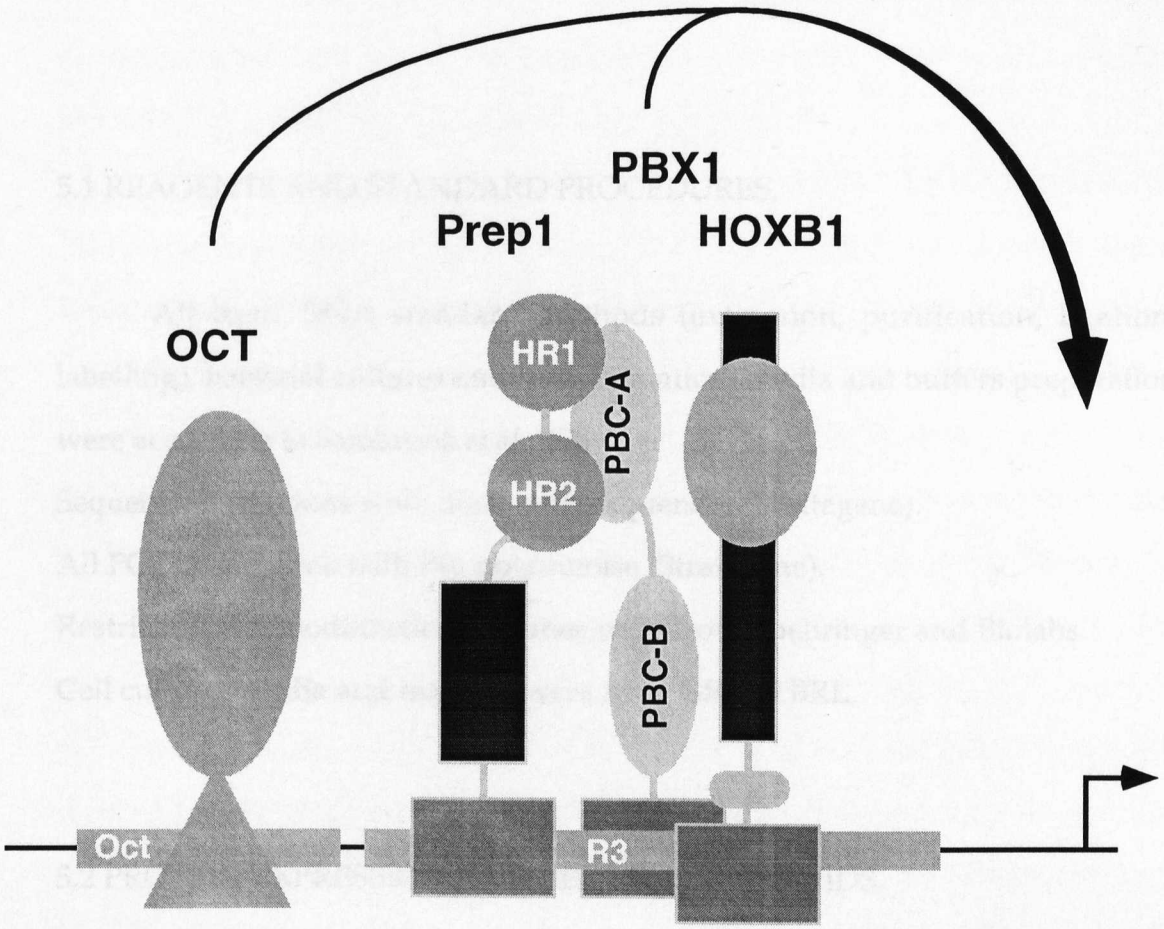


Figure 4.1: A model for the ARE element in a transcriptional active configuration. Here is represented the 3' region of the enhancer, corresponding to the ΔR1+R2 fragment, and including only the octamer and the R3 sites. Pbx is interacting simultaneosly with two proteins: HoxB1, via its homeodomain, and Prep1, via its N-terminal region PBC-A. An octamer binding protein (OCT) sinergyzes with the trimeric complex to activate transcription.

Chapter 5

MATERIALS AND METHODS

5.1 REAGENTS AND STANDARD PROCEDURES.

All basic DNA standard methods (extraction, purification, ligation, labelling), bacterial cultures and transformation, media and buffers preparation were according to Sambrook *et al.*, 1989.

Sequencing reactions were done with Sequenase (Stratagene).

All PCRs were done with Pfu polymerase (Stratagene).

Restriction and modification enzymes were from Boehringer and Biolabs.

Cell cultures media and reagents were from GIBCO BRL.

5.2 PROTEIN EXPRESSION AND REPORTER PLASMIDS.

All expression constructs are derivatives of the SV40 promoter driven expression vector pSG5 (Green *et al.*, 1988). Hox expression vectors were previously described (Arcioni *et al.*, 1992; Guazzi *et al.*, 1994; Zappavigna *et al.*, 1991; Zappavigna *et al.*, 1994), with the exception of pSGHOXA1 and pSGHOXC6. pSGHOXA1 was generated by ligating a Bam HI-Xba I Klenow-filled *Hoxa-1* cDNA coding sequence obtained from the p*Hoxa-1*(HD$^+$) plasmid (Phelan *et al.*, 1995) into the Bam HI-Bgl II Klenow-filled sites of pSG5. pSGHOXC6 was obtained cloning a PCR amplified Bam HI insert containing

the full length HOXC6 coding sequence from the pCT-H3C plasmid (Arcioni *et al.*, 1992) into the Bam HI site of pSG5.

pSGPbx1a was generated ligating an Hind III filled-Eco RI filled cDNA fragment encompassing the complete coding sequence of Pbx1a (Van Dijk *et al.*, 1993) into the filled Bam HI site of pSG5. pSGPbx1b was reconstructed by substituting the E1A N-terminal region of E1APbx1b cDNA at an unique Nco I site with the N-terminal region of Pbx1a

Pbx1a deletion mutants (pSGpbx1a Δ1-140, Δ1-230, Δ296-430 and Δ283-285) were generated by ligating Bam HI PCR-amplified and deleted cDNA fragments in the Bam HI site of pSG5. To generate pSGPbx1a Δ1-140 and Δ1-230 an ATG start codon was introduced in frame upstream of amino acids 141 and 231, respectively. To generate pSGPbx1a Δ296-430, a TGA stop codon was added downstream of amino acid 295. pSGpbx1b Δ1-230 was generated using the same PCR primer as for Pbx1aΔ1-230 and cloning into the Eco RI-Bam HI sites of pSG5. pSGB1 ΔWM and W177A were obtained by PCR mutagenesis, covering the full length HOXB1, as Bam HI inserts into pSG5. In pSGB1ΔWM, amino acids W177 and M178 of the HOXB1 "YPWM" motif were deleted and in pSGB1W177A amino acid W177 was replaced by an alanine. pSGB1Δ236-274 was obtained by removing the Sac I-Pvu II fragment from pSGHOXB1 and re-ligating the vector after blunting of the Sac I site with T4 DNA polymerase. This generates a deletion in the HOXB1 cDNA extending from the C-terminal region of the homeodomain (helix 3/4) to part of the adjacent C-terminus of the protein, without any amino acid substitution.

pSGB3/B1HN was generated by substituting the "YPWM"/homeodomain N-terminal region of HOXB3 (amino acids 130-201) with the analogous region of HOXB1 (amino acids 175-211) while pSGB3/B1N was generated by replacing the homeodomain N-terminal region of HOXB3 (amino acids 184-201) with the analogous region of HOXB1 (amino acids 195-

211). The mutated cDNAs were obtained by PCR amplification and cloned as Bam HI inserts into pSG5. pSGB1HD was generated by PCR amplification of a region comprising amino acids 155-286 and introducing a methionine residue in position 154.

pSGGAL1-147 was obtained cloning a Bam HI-Bgl II insert containing the DNA binding domain of yeast GAL4 (amino acids 1 to 147) into pSG5. pGAL1-147 pbx1CT was generated by PCR amplifying a region comprising amino acids 348-430 and cloning in frame with the GAL4 1-147 protein at the Eco RI site of the pGAL1-147 vector. The correctness of all cloned PCR product was verified by DNA sequencing and the expression of all proteins was tested in a rabbit reticulocyte system.

The luciferase reporter construct pML is a pXP2 based vector (Nordeen, 1988) containing the Adenovirus Major Late basal promoter (from -65 to +30). pAdMLARE contains the Ava I-Hae II r4 enhancer of *Hoxb-1* (Pöpperl *et al.*, 1995) cloned as a PCR amplified Hind III-Xho I fragment into pML. In pAdMLAREmR3, the Pbx consensus site of repeat 3 in the r4 enhancer TGATGGAT was changed to TGTCGACT. pAdMLR3 contains a trimer of repeat 3 of the *Hoxb-1* ARE enhancer cloned as a Bam HI-Hind III fragment into pML. The same trimer of repeat 3 was cloned into the Bam HI site of the pT81luc luciferase reporter vector (Nordeen, 1988) containing the Thymidine Kinase basal promoter. The sequence of the 30 bp oligo used to generate the trimer is 5'-GATCCGGGGGGTGATGGATGGGCGCTGGGA-3'. pAdMLARE Δ1-52, ΔR1 and ΔR1+R2 have been obtained by cloning into pML PCR amplified Hind III-Xho I fragments containing the *Hoxb-1* r4 enhancer from nucleotide 53 to 140, 64 to 140 and 82 to 140, respectively (see Fig. 1.6 for ARE sequence). In pAdMLoctm, the octamer-like sequence ATGCTAAT in the r4 enhancer was changed to GACTGCCT. The pTUASluc-GAL4 reporter construct was described in Zappavigna *et al.*, 1996.

5.3 CELL CULTURES AND TRANSFECTIONS.

COS7 and P19 cells were maintained in Dulbecco's Medium (DMEM) and MEM-α medium, respectively, supplemented with 10% fetal calf serum (GIBCO), 100 I.U./ml of penicillin and 100 µg/ml streptomycin. Cells were transfected by calcium-phosphate precipitation (Di Nocera and Dawid, 1983). A precipitate containing calcium phosphate and DNA is formed by slowly mixing a Hepes-buffered saline solution (prepared by mixing 16.4 g NaCl, 11.9 g Hepes acid, 0.21 g Na_2PO4 in 1 litre of H_2O; pH adjusted to 7.05-7.12.) with a solution containing $CaCl_2$ 125 mM and DNA purified in a CsCl gradient. The precipitate is allowed to stand for 30 min at room temperature and then gently distributed on cell plates which were previously fed with fresh medium. In a typical transfection experiment, 8 µg of reporter plasmid, 4-8 µg of expression construct and 0.2 µg of pCMV-ß-gal (Clontech) as an internal control, were used per 10 cm dish. 48-60 hrs after transfection cells were washed and lysed directly on the plate with a solution containing Triton 1%, Glycil-Glycine pH7.8 25 mM; $MgSO_4$ 15 mM; EGTA 4 mM DTT 1 mM. Extracts were collected, centrifuged to clear the supernatant and assayed for luciferase and ß-galactosidase expression as described in Sambrook *et al.*, 1989.

5.4 PROTEIN PRODUCTION AND CELLULAR EXTRACTS.

Pbx and HOX proteins were produced *in vitro* from the corresponding pSG5-derived expression vectors using a T7 polymerase based coupled transcription/translation reticulocyte lysate system (Promega, Madison, Wisconsin) according to recommended conditions. HOX and Pbx proteins were translated separately in the presence of ^{35}S methionine and quantitated after

SDS-PAGE using a PhosphorImager (Molecular Dynamics). The amounts of proteins were normalised for the methionine content of each protein.

For total cellular extracts, transfected or non-transfected cells were collected from confluent plates, washed with PBS (phosphate buffered saline), pelleted, freezed with liquid nitrogen and lysed by resuspending in 5 volumes of Extraction Buffer (10 mM Hepes pH 7.9, O.4 M NaCl, 0.1 mM EDTA, 0.5 mM DTT, 5% Glycerol, O.5 mM PMSF, 1% Trasylol). The lysate is then centrifuged for 30 min at 34000 rpm in a Beckman Ti50 rotor and the supernatant stored in aliquots at -80 °C.

For total embryo extracts, about 50 embryos (9.5 dpc) were collected from the uterus of pregnant mice, washed with PBS and then lysed as described above.

5.5 ELECTROPHORETIC MOBILITY SHIFT ASSAYS (EMSA).

Gel-retardation analysis was performed by pre-incubating the *in vitro* synthesized proteins (2-5 µl) or total cellular extracts (10 µg) for 30 minutes on ice in 20 µl of binding buffer (75 mM NaCl, 6% glycerol, 10 mM Tris-HCl pH 7.6, 1 mM EDTA), together with 2 µl (0.5 ng, 5 x 10^4 cpm) of [^{32}P]-labelled probe. Competition were performed by adding a 200 fold excess of unlabelled fragments or 2µl of antiserum before adding the probe. The incubation mixture was resolved by electrophoresis on a 6% polyacrylamide gel in 0.25 x TBE at 10 V/cm. Gels were dried and exposed to a Kodak X-AR film at -70 °C.

The ΔR1+R2 probe is a PCR amplified fragment corresponding to nucleotides 82-140 of the b-1ARE enhancer sequence (Fig.1.6). The R3 oligonucleotide probe sequence is:

5'-GATCCGGGGGGTGATGGATGGGCGCTGGGA-3'.

α-Oct1 antibody was purchased by Santa Cruz Biotechnology. α-Oct3 antibody was kindly provided by H. Shöler.

5.6 NUCLEAR CELLULAR EXTRACTS AND DNASE I FOOTPRINT.

Nuclear extracts were according to Dignam *et al.*, 1983, with minor modifications: P19 and Cos7 cells were washed with PBS, collected from plates by scraping, and harvested by centrifugation. The pellet was resuspended in five packed cell volumes of buffer A (Hepes pH 7.9 10 mM; $MgCl_2$ 1.5 mM; KCl 10 mM, DTT 0.5 mM, Trasylol 5μg/ml, PMSF 0.5 mM) and allowed to stand on ice for 10 min. Cells were then pelleted, resuspended in two volumes of buffer A and transferred to a Dounce homogenizer to lyse the cytoplasm. Nuclei were pelleted, washed with buffer A, and resuspended in 0.5 volumes of buffer C (Hepes pH 7.9 20 mM; Glycerol 25%; NaCl 420 mM; $MgCl_2$ 1.5 mM; EDTA 0.2 mM; DTT 0.5 mM, Trasylol 5μg/ml, PMSF 0.5 mM) in a Dounce homogenizer. The resulting suspension was stirred gently for 30 min and then centrifuged 30 min at 25.000 g. The resulting clear supernatant was dialyzed against 50 volumes of buffer D (Hepes pH 7.9 20 mM; Glycerol 20%; KCl 0.1 mM; EDTA 0.2 mM; DTT 0.5 mM, Trasylol 2μg/ml, PMSF 0.5 mM) for five hours. The dialysate was centrifuged at 25.000 g for 20 minutes and the supernatant was frozen in aliquots in liquid nitrogen and stored at -80 °C.

For footprint reactions, 20.000 cpm of end labelled AREΔR1+R2 fragment corresponding to the 81-140 region of the b-1ARE enhancer were incubated at room temperature with 10 or 20μg of cellular nuclear extract in a buffer containing 1μg of poly d(I-C); KCl 50 mM; $MgCl_2$ 2 mM; Glycerol 20%; Tris-HCl pH7.9 20 mM; DTT 1 mM in a total volume of 50 μl. Samples were digested for 90 sec with 300 ng of DNase I (50 ng for control DNA without the extract).

The reaction was stopped by adding 100 µl of Stop solution (NaCl 0.2 M; EDTA 30 mM; SDS 1%; yeast t-RNA 100 µg/ml). The digested DNA was then extracted with phenol-chloroform, ethanol precipitated, resuspended in a formamide loading dye, denatured for 5 min at 95 °C and loaded onto a 6% acrylamide - 50% urea sequencing gel. Gel were dried and exposed overnight at -80 °C to Kodak X-AR5 films.

G+A sequencing reactions were performed as follows: the probe DNA fragment (200.000 cpm in 10 µl reaction volume) together with 1µg of carrier DNA was incubated with 25 µl of pure formic acid for 5 min at room temperature. The reaction was stopped by adding 200 µl of HZ Stop mix (Na acetate pH 7.5 0.3 M; EDTA 0.1 mM; yeast t-RNA 25 mg/ml). The sample was then precipitated with ethanol, resuspended in 100 µl of piperidine 1 M and incubated at 90 °C for 30 min. Piperidine was removed by lyophilizations and the sample resuspended in formamide loading dye. About 10.000 cpm were used per each lane.

REFERENCES

Arcioni, L., Simeone, A., Guazzi, S., Zappavigna, V., Boncinelli, E. and Mavilio, F. (1992) The upstream region of the human homeobox gene HOX3D is a target for regulation by retinoic acid and HOX homeoproteins. *EMBO J.*, **11**, 265-277.

Bally-Cuif, L. and Boncinelli, E. (1996) Transcription factors and head formation in vertebrates. *BioEsssays*, **19**, 127-135.

Barrow, J. R. and Capecchi M. R. (1996) Targeted disruption of the *Hoxb-2* locus in mice interferes with expression of *Hoxb-1* and *Hoxb-4*. *Development*, **122**, 3817-3828.

Benson, G. V., Lim, H., Paria, B.C., Stokata, I., Dey, S. K. and Maas, R.L. (1996) Mechanisms of reduced fertility in *Hoxa-10* mutant mice: uterine homeosis and loss of maternal *Hoxa-10* expression. *Development*, **122**, 2687-2696.

Berthelsen, J., Vandekerkhove, J. and Blasi F. (1996) Purification and characterization of UEF3, a novel factor involved in the regulation of the urokinase and other AP-1 controlled promoters. *J. Biol. Chem.*, **271**, 3822-3830.

Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F. and Blasi, F. (1998) The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity. *EMBO J.*, **17**, 1434-1445.

Billeter, M., Qian, Y., Otting, G., Müller, M., Gehring, W., and Wüthrich, K. (1993) Determination of the NMR solution structure of an Antennapedia homeodomain-DNA complex. *J. Mol. Biol.*, **234**, 1094-1097.

Bürglin, T. R. and Ruvkun, G. (1992) New motif in PBX genes. *Nat. Genet.*, **1**, 319-320.

Bürglin, T. R. (1997) Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Res.*, **25**, 4173-4180.

Carpenter, E. M., Goddard, J. M., Chisaka, O., Manley, N. R. and Capecchi, M. R. (1993) Loss of *Hoxa-1* (*Hox-1.6*) function results in the reorganization of the murine hindbrain. *Development*, **118**, 1063-1075.

Casares, F. and Mann, R. S. (1998) Control of antennal versus leg development in Drosophila [see comments]. *Nature*, **392**, 723-6.

Chan, S.-K. and Mann, R. S. (1993) The segment identity functions of Ultrabithorax are contained within its homeodomain and carboxy-terminal sequences. *Genes Dev.*, **7**, 796-811.

Chan, S.-K., Jaffe, L., Capovilla, M., Botas, J. and Mann, R. S. (1994) The DNA binding specificity of ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. *Cell*, **78**, 603-615.

Chan, S.-K. and Mann, R. S. (1996) A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proc. Natl. Acad. Sci. USA*, **93**, 5223-5228.

Chan, S.-K., Pöpperl, H., Krumlauf, R. and Mann, R. S. (1996) An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J.*, **15**, 2476-2487.

Chan, S.-K., Ryoo, H.-D., Gould, A., Krumlauf, R. and Mann, R. (1997). Switching the in vivo specificity of a minimal Hox responsive element. *Development*, **124**, 2007-2014.

Chang, C.-P., Shen, W.-F., Rozenfeld, S., Lawrence, H. J., Largman, C. and Cleary, M. (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev.*, **9**, 663-674.

Chang, C.-P., Brocchieri, L., Shen, W.-F., Largman, C. and Cleary, M. (1996) Pbx modulation of Hox homeodomain Amino-terminal arms establishes different DNA-binding specificities across the *Hox* locus. *Mol. Cell. Biol.*, **16**, 1734-1745.

Davis, A. P. and Capecchi, M. R. (1996) A mutational analysis of the 5' HoxD genes: dissection of genetic interactions during limb development in the mouse. *Development*, **122**, 1175-1185.

De Cesare, D., Palazzolo, M., Berthelsen, J. and Blasi, F. (1997) Characterization of UEF-4, a DNA-binding protein required for transcriptional synergism between two AP-1 sites in the human urokinase enhancer. *J. Biol. Chem.*, **272**, 23921-23929.

Dedera, D. A., Waller, E. K., Lebrun, D.P., Sen-Majumdar, A., Stevens, M. E., Barsh, G. S. and Cleary, M. L. (1993) Chimeric homeobox gene E2A-Pbx1 induces proliferation, apoptosis, and malignant lymphomas in transgenic mice. *Cell*, **74**, 833-
843.

Dignam, J. R., Lebovitz, R. M. and Roeder, R. G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucl. Acids Res.* **11**, 1475-1489.

Di Nocera, P. P. and Dawid, I. B. (1983) Transient expression of genes introduced into cultured cells of *Drosophila*. *Proc. Natl. Acad. Sci. USA*, **80**, 7095-7098.

Dollé, P., Lufkin, T., Krumlauf, R. Mark, M., Doubule, D. and Chambon, P. (1993) Local alterations of Krox-20 and Hox gene expression in the hindbrain of Hoxa-1 (Hox-1.6) homozygote null mutant embryos. *Proc. Natl. Acad. Sci. USA* , **90**, 7666-7670.

Douboule D. and Morata G. (1994) Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.*, **10**, 358-364.

Favier, B., Rijli F.M., Fromental-Ramain, C., Fraulob, V., Chambon, P. and Dollé, P. (1996) Functional cooperation between the non paralogous genes *Hoxa-10* and *Hoxd-11* in the developing forelimb and axial skeleton. *Development*, **122**, 449-460.

Favier, B. and Dollé P. (1997) Developmental functions of mammalian Hox genes. *Mol. Hum. Reprod.*, **3**, 115-131.

Fraser, S., Keynes, R. and Lumsden, A. (1990) Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature*, **344**, 431-435.

Fromental-Ramain, C., Warot, X., Lakkaraju, S., Favier, B., Haack, H., Birling, C., Dierich, A., Dollé, P. and Chambon, P. (1996) Specific and redundant functions of the paralogous *Hoxa-9* and *Hoxd-9* genes in forelimb and axial skeleton patterning. *Development*, **122**, 461-472.

Furukubo-Tokunaga, K., Flister, S. and Gehring, W. J. (1993) Functional specificity of the Antennapedia homeodomain. *Proc. Natl. Acad. Sci. USA*, **90**, 6360-6364.

Gehring W. J. (1994). A history of the Homeobox. In: Douboule D. (ed) Guidebook to the Homeobox Genes. Oxford University Press, Oxford.

Gehring, W. J., Quian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994) Homeodomain-DNA recognition. *Cell*, **78**, 211-223.

Gendron-Maguire, M., Mallo, M., Zhang, M. and Gridley, T. (1993) *Hoxa-2* mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell*, **75**, 1317-1331.

Goddard, J. M., Rossel, M., Manley, N. R. and Capecchi M. R. (1993) Loss of *Hoxa-1* (*Hox-1.6*) function results in the reorganization of the murine hindbrain. *Development*, **118**, 1063-1075.

Gonzales-Crespo, S. and Morata, G. (1995) Control of *Drosophila* adult pattern by *extradenticle*. *Development*, **121**, 2117-2125.

Green, S., Issemann, I. and Sheer, E. (1988) A versatile eucaryotic expression vector for protein engineering. *Nucl. Acids Res.*, **16**, 369-373.

Grieder, N. C., Marty, T., Ryoo, H. D., Mann R. S. and Affolter, M. (1997) Synergistic activation of a Drosophila enhancer by HOM/EXD and DPP signaling. *EMBO J.*, **16**, 7402-7410.

Gross, C. T. and McGinnis, W. (1996) DEAF-1, a novel protein that binds an essential region in a Deformed response element. *EMBO J.*, **15**, 1961-1970.

Grueneberg, D.A., Natesan, S., Alexandre, C. and Gilman, M. Z. (1992) Human and *Drosophila* homeodomain proteins that enhance the DNA-binding activity of serum response factor. *Science*, **257**, 1089-1095.

Guazzi, S., Lonigro, R., Pintonello, L., Boncinelli, E., Di Lauro, R. and Mavilio, F. (1994) The thyroid transcriptional factor-1 gene is a candidate target for regulation by Hox proteins. *EMBO J.*, **13**, 3339-3347.

Hauptmann, G. and Gerster, T. (1995) Pou-2 - a zebrafish gene active during cleavage stages and in the early hindbrain. *Mech. Dev.*, **51**, 127-138.

Herr, W. and Cleary, M. A. (1995) The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes Dev.*, **9**, 1679-1693.

Hsieh-Li, H.M., Witte, D.P., Weistain, M., Branford, W., Li, H., Small, K. and Potter S. S. (1995) *Hoxa-11* structure, extensive antisense transcription, and function in male and female fertility. *Development*, **121**, 1373-1385.

Johnson, F. B., Parker, E. and Krasnow, M. (1995) Extradenticle protein is a selective cofactor for the *Drosophila* homeotics: Role of the homeodomain and YPWM amino acid motif in the interaction. *Proc. Natl. Acad. Sci. USA*, **92**, 739-743.

Jones, F. S., Chalepakis, G., Gruss, P. and Edelman, G. M. (1992) Activation of the cytotactin promoter by the homeobox-containing gene Evx-1. *Proc. Natl. Acad. Sci. USA*, **89**, 2091-2095.

Jones, F. S., Holst, B. D., Minowa, O., De Robertis, E. M. and Edelman, G. M. (1993) Binding and transcriptional activation of the promoter for the neural cell adhesion molecule by *HoxC6 (Hox-3.3)*. *Proc. Natl. Acad. Sci. USA*, **90**, 6557-6561.

Kamps, M. P., Murre, C., Sun, X.-H. and Baltimore, D. (1990) A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in Pre-B-ALL. *Cell,*, **60**, 547-555.

Kamps, M. P., Wright, D. and Lu, Q. (1996) DNA-binding by oncoprotein E2A-Pbx1 is important for blocking differentiation but dispensable for fibroblast transformation. *Oncogene*, **12**, 19-30.

Kessel M., Balling R. and Gruss P. (1990) Variation of cervical vertebrae after expression of a *Hox-1:1* transgene in mice. *Cell* , **61**, 301-308.

Kessel M. and Gruss P. (1991) Homeotic transformation of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell*, **67**, 89-104.

Keynes, R. and Krumlauf, R. (1994) HOX genes and the regionalization of the nervous system. *Ann. Rev. Neurosci.*, **17**, 109-132.

Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. and Pabo, C. O. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell*, **63**, 579-590.

Knoepfler, P. and Kamps, M. (1995) The pentapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding by Pbx1. *Mol. Cell. Biol.*, **15**, 5811-5819.

Knoepfler, P., Lu, Q. and Kamps, M. (1996) Pbx1-Hox heterodimers bind DNA on inseparable half-sites that permit intrinsic DNA binding specificity of the Hox partner at nucleotides 3' to a TAAT motif. *Nucleic Acid sRes.*, **24**, 2288-2294.

Kondo, T., Dollé, P., Zàkàny J. and Doubule, D. (1996) Function of posterior HoxD genes in the morphogenesis of the anal sphincter. *Development*, **122**, 2651-2659.

Köntges, G. and Lumsden, A. (1996) Rhombencefalic neural crest segmantation is preserved throughout craniofacial ontogeny. *Development*, **122**, 3229-3242.

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

Kurant, E., Pai, C. Y., Sharf, R., Halachmi, N., Sun, Y. H. and Salzberg, A. (1998) Dorsotonals/homothorax, the Drosophila homologue of meis1, interacts with extradenticle in patterning of the embryonic PNS. *Development*, **125**, 1037-1048.

Lai, J. S., Cleary, M. A. and Herr, W. (1992) A single aminoacid exchange transfers VP16-induced positive control from the Oct-1 to the Oct-2 homeodomain. *Genes Dev.*, **6**, 205-2065.

Lawrence P. A. and Morata G. (1994) Homeobox genes: their function in *Drosophila* segmentation and pattern formation. *Cell*, **78**, 181-189.

Le Mouellic, H., Lallemand, Y. and Brulet, P. (1992) Homeosis in the mouse induced by a null mutation in the *Hox 3:1* gene. *Cell*, **69**, 251-269.

Lewis, E. B. (1978) A gene complex controlling segmentation in *Drosophila*. *Nature*, **276**, 565-570.

Lin, L. and McGinnis, W. (1992) Mapping functional specificity in the Dfd and Ubx homeodomains. *Genes Dev.*, **6**, 1071-1081.

Lu, Q., Knoepfler, P., Scheele, J., Wright, D. and Kamps, M. (1995) Pbx1 and E2A-Pbx1 bind the DNA motif ATCAATCAA cooperatively with the products of multiple murine Hox genes, some of which are themselves oncogenes. *Mol. Cell. Biol.*, **15**, 3786-3795.

Lu, Q. and Kamps, M. P. (1996a) Selective repression of transcriptional activators by Pbx1 does not require the homeodomain. *Proc. Natl. Acad. Sci. USA*, **93**, 470-474.

Lu, Q. and Kamps, M. P. (1996b) Structural determinants within Pbx1 that mediate cooperative DNA binding with pentapeptide-containing Hox proteins: proposal for a model of a Pbx1-Hox-DNA complex. *Mol. Cell. Biol.*, **16**, 1632-1640.

Lufkin T., Mark M., Hart C. P., Dollé P., LeMeur M. and Chambon P. (1992) Homeotic transformation of the occipital bones of the skull by ectopic expression of a homeobox gene. *Nature*, **359**, 835-841.

Lumsden, A. and Krumlauf, R. (1996) Patterning the vertebrate neuraxis. *Science*, **274**, 1109-1115.

Lutz, B., Lu, H. C., Eichele, G., Miller, D. and Kaufmann, T.C. (1996) Rescue of *Drosophila labial* null mutant by the chicken ortholog *Hoxb-1* demonstrates that the function of Hox genes is phylogenetically conserved. *Genes Dev.*, **10**, 176-184.

Maconochie, M., Nonchev, S., Morrison, A. and Krumlauf, R. (1996) Paralogous Hox genes: function and regulation. *Annu. Rev. Genet.*, **30**, 529-556.

Maconochie, M., Nonchev, S., Studer, M., Chan, S.-K., Pöpperl, H., Sham, M. H., Mann, R. and Krumlauf, R. (1997) Cross regulation in the mouse HoxB complex: the expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1. *Genes. Dev.*, **11**, 1885-1895.

Manak, J. R. and Scott, M. P. (1993) Able assistants for homeodomain proteins. *Current Biology*, **3**, 318-320.

Manak, J. R. and Scott, M. P. (1994) A class act: conservation of homeodomain protein functions. *Development*, Suppl., 61-71.

Mann, R. (1995) The specificity of homeotic gene function. *Bioessays*, **17**, 855-863.

Mann, R. S. and Chan, S.-K. (1996) Extra specificity from *extradenticle*: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet.*, **12**, 258-262.

Mann, R. S. and Affolter, M (1998) Hox proteins meet more partners. *Curr. Op. Gen. Dev.*, **8**, 423-429.

Mark, M., Lufkin, T., Vonesch, J.-L., Ruberte, E., Oliver, J.-C., Dolle, P., Gorry, P., Lumsden, A. and Chambon, P. (1993) Two rhombomeres are altered in *Hoxa-1* mutant mice. *Development*, **119**, 319-338.

Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Zappavigna, V., Acampora, D., Poiana, G., Russo, G., Peschle, C. and Boncinelli, E. (1986) Differential and stage related expression in embryonic tissues of a new human homeobox gene. *Nature*, **324**, 664-668.

McGinnis W. and Krumlauf R. (1992) Homeobox genes and axial patterning. *Cell*, **68**, 283-302.

Monica, K., Galili, N., Nourse, J., Saltman, D. and Cleary, M. L. (1991) *PBX2* and *PBX3*, new homeobox genes with extensive homology to the human proto-oncogene *PBX1*. *Mol. Cell. Biol.*, **11**, 6149-6157.

Nakamura, T., Jenkins, N. A. and Copeland, N. G. (1996) Identification of a new family of Pbx-related homeobox genes. *Oncogene*, **13**, 2235-2242.

Nicosia, A., Monaci, P., Tomei, L., DeFrancesco, R., Nuzzo, M., Stunnenberg, H. and Cortese, R. (1990) A myosin-like dimerization helix and an extra-large homeodomain are essential elements of the tripartite DNA-binding structure of LFB-1. *Cell*, **61**, 1225-1236.

Nordeen, S. K. (1988) Luciferase reporter gene vectors for analysis of promoters and enhancers. *Biotechniques*, **6**, 454-457.

Nourse, J., Melletntin, J., Galili, N., Wilkinson, J., Stanbridge, E., Smith, S. and Cleary, M. (1990) Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell*, **60**, 535-545.

Pai, C. Y., Kuo, T. S., Jaw, T. J., Kurant, E., Chen, C. T., Bessarab, D. A. Salzberg, A., and Sun, Y. H. (1998) The Homothorax homeoprotein activates the nuclear localization of another homeoprotein, extradenticle, and suppresses eye development in Drosophila. *Genes Dev.*, **12**, 435-446.

Peers, B., Sharma, S., Johnson, T., Kamps, M. and Montminy, M. (1995) The pancreatic islet factor STF-1 binds cooperatively with Pbx to a regulatory element in the somatostatin promoter: importance of the FPWMK motif and of the homeodomain. *Mol. Cell. Biol.*, **15**, 7091-7097.

Peifer, M. and Wieschaus, E. (1990) Mutations in the *Drosophila* gene *extradenticle* affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev.*, **4**, 1209-1223.

Peltenburg, L. T. C. and Murre, C. (1996) Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx. *EMBO J.*, **15**, 3385-3393.

Phelan, M. L., Rambaldi, I. and Featherstone, M. S. (1995) Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. *Mol. Cell. Biol.*, **15**, 3989-3997.

Pöpperl, H., Bienz, M., Studer, M., Chang, S.-K., Aparicio, S., Brenner, S., Mann, R. and Krumlauf, R. (1995) Segmental expression of *Hoxb-1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/Pbx*. *Cell*, **81**, 1031-1042.

Pöpperl, H. and Featherstone, M. S. (1992) An autoregulatory element of the murine *Hox-4.2* gene. *EMBO J.*, **11**, 3673-3680.

Rauskolb, C., Peifer, M. and Wieschaus, E. (1993) *extradenticle*, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene *pbx1*. *Cell*, **74**, 1101-1112.

Rauskolb, C. and Wieschaus, E. (1994) Coordinate regulation of downstream genes by *extradenticle* and the homeotic selector proteins. *EMBO J.*, **13**, 3561-3569.

Rijli, F., Mark, M., Lakkaraju, S., Dierich, A., Dollé, P. and Chambon, P. (1993) A homeotic transformation is generated in the rostral branchial region of the head by dirsuption of *Hoxa-2*, which acts as a selector gene. *Cell*, **75**, 1333-1349.

Rudnicki, M. A. and McBurney, M. W. (1987) Cell culture methods and induction of differentiation of embryonal carcinoma cell lines. In: Teratocarcinomas and embryonic stem cells. A practical approach. IRL Press.

Ryan, A. K. and Rosenfeld, M. G. (1997) POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev.*, **11**, 1207-1225.

Sambrook, J., Fritsch, E. F. and Maniatis T. (1989) Molecular cloning - A laboratory manual. Second edition. CSH Laboratory Press; Cold Spring Harbor, N.Y.

Sauer, F., Hansen, S.K. and Tijan, R. (1995) Multiple TAFIIs directing synergistic activation of transcription. *Science*, **270**, 1783-1788.

Shen, W.-F., Rozenfeld, S., Lawrence, H. J. and Largman, C. (1997) The Abd-B like Hox homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on novel DNA target. *J. Biol. Chem.*, **272**, 8198-8206.

Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E. and Mavilio, F. (1990) Sequential activation of *Hox2* homeobox genes by retinoic acid in human embryonal carcinoma cells. *Nature*, **346**, 763-766.

Smith, D. L. and Johnson, A. D. (1992) A molecular mechanism for combinatorial control in yeast: MCM1 protein sets the spacing and orientation of the homeodomains of an α2 dimer. *Cell*, **68**, 133-142.

Stein S., Fritsch R., Lemaire L. and Kessel M. (1996) Checklist: vertebrate homeobox genes. *Mech. Dev.*, **55**, 91-108.

Struhl, G. (1981) A homeotic mutation transforming leg to antenna in *Drosophila.. Nature* , **292**, 635-638.

Studer, M., Lumsden, A., Ariza-McNaughton, L., Bradley, A. and Krumlauf, R. (1996) Altered segmental identity and abnormal migration of motor neurons in mice lacking *Hoxb-1. Nature*, **384**, 630-634.

Sun, B., Hursh, D. A., Jackson, D. and Beachy, P. A. (1995) Ultrabithorax protein is necessary but not sufficient for full activation of *decapentaplegic* epression in the visceral mesoderm. *EMBO J.*, **14**, 520-535.

Treisman, J., Gönczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989) A single aminoacid can determine the DNA binding specificity of homeodomain proteins. *Cell*, **59**, 553-562.

Van Dijk, M. A. and Murre, C. (1994) *extradenticle* raises the DNA binding specificity of homeotic selector gene products. *Cell*, **78**, 617-624.

Van Dijk, M. A., Voorhoeve, P. M. and Murre, C. (1993) Pbx1 is converted into a transcriptional activator upon acquiring the N-terminal region of E2A in pre-B-cell acute lymphoblastoid leukaemia. *Proc. Natl. Acad. Sci. USA*, **90**, 6061-6065.

Vershon, A. K., Jin, Y. S. and Johnson, A. D. (1995) A homeodomain protein lacking specific side chains of helix 3 can still bind DNA and direct transcriptional repression. *Genes Dev.*, **9**, 182-192.

Wieschaus, E. and Noell, E. F. (1986) Specificity of embryonic lethal mutations in *Drosophila* analysed in germ line clones. *Roux's Arch. Dev. Biol.*, **195**, 63-73.

Viganò, A. and Staudt, L. M. (1996) Transcriptional activation by Oct-3: evidence for a specific role of the POU-specific domain in mediating functional interaction with Oct-1. *Nucleic Acids Res.*, **24**, 2112-2118.

Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. and Pabo, C. O. (1991) Crystal structure of MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell*, **67**, 517-528.

Zàkàny, J. and Douboule, D. (1996) Synpolydactyly in mice with a targeted deficiency in the HoxD complex. *Nature*, **384**, 69-71.

Zappavigna, V., Falciola, L., Helmer Citterich, M., Mavilio, F. and Bianchi, E. M. (1996) HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. *EMBO J.*, **15**, 4981-4991.

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**, 4177-4187.

Zappavigna, V., Sartori, D. and Mavilio, F. (1994) Specificity of HOX protein function depends on DNA-protein and protein-protein interactions, both mediated by the homeo domain. *Genes Dev.*, **8**, 732-744.

Zappavigna, V., Falciola, L., Helmer Citterich, M., Mavilio, F. and Bianchi, M. E. (1996) HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. *EMBO J.*, **15**, 4981-4991.

Zeng, W., Andrew, D. J., Mathies, L. D., Horner, M. A. and Scott, M. P. (1993) Ectopic expression and function of the Antp and Scr homeotic genes: the N-terminus of the homeodomain is critical to functional specificity. *Development*, **118**, 339-352.

Zhang, M., Kim, H.-J., Marshall, H., Gendron-Maguire, M., Lucas, A. D., Baron, A., Gudas, L. J., Gridley, T., Krumlauf, R. and Grippo, J. F. (1994) Ectopic *Hoxa-1* induces rhombomere transformation in mouse hindbrain. *Development*, **120**, 2431-2442.