ARTICLE
Journal of Cellular Biochemistry 110:484–496 (2010)

Journal of Cellular Biochemistry

Pentapeptide Insertion Mutagenesis of the Hoxa1 Protein: Mapping of Transcription Activation and DNA-Binding Regulatory Domains

Barbara Lambert,[1] Julie Vandeputte,[1] Pierre-Marie Desmet,[1] Bernard Hallet,[2] Sophie Remacle,[1] and René Rezsohazy[1*]

[1] *Unit of Veterinary Sciences, Life Sciences Institute (ISV), Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium*

[2] *Unit of Genetics, Life Sciences Institute (ISV), Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium*

ABSTRACT

The mode of action of Hoxa1, like that of most Hox proteins, remains poorly characterized. In an effort to identify functional determinants contributing to the activity of Hoxa1 as a transcription factor, we generated 18 pentapeptide insertion mutants of the Hoxa1 protein and we assayed them in transfected cells for their activity on target enhancers from the *EphA2* and *Hoxb1* genes known to respond to Hoxa1 in the developing hindbrain. Only four mutants displayed a complete loss-of-function. Three of them contained an insertion in the homeodomain of Hoxa1, whereas the fourth loss-of-function mutant harbored an insertion in the very N-terminal end of the protein. Transcription activation assays in yeast further revealed that the integrity of both the N-terminal end and homeodomain is required for Hoxa1-mediated transcriptional activation. Furthermore, an insertion in the serine–threonine–proline rich C-terminal extremity of Hoxa1 induced an increase in activity in mammalian cells as well as in the yeast assay. The C-terminal extremity thus modulates the transcriptional activation capacity of the protein. Finally, electrophoretic mobility shift assays revealed that the N-terminal extremity of the protein also exerts a modulatory influence on DNA binding by Hoxa1–Pbx1a heterodimers. J. Cell. Biochem. 110: 484–496, 2010.

KEY WORDS: HOMEODOMAIN; TRANSCRIPTION FACTOR; Hoxa1; Pbx; Hoxb1; EphA2; PENTAPEPTIDE SCANNING MUTAGENESIS

Hox genes encode a highly conserved family of transcription factors characterized by a common sequence element of 180 bp, the homeobox. Mammalian genomes share 39 *Hox* genes that are clustered in four tight chromosomal loci. Based on their sequence similarity, these genes have been classified in 13 paralog groups. The *Hox* genes play critical roles in governing developmental programs leading to embryo patterning, morphogenesis and organogenesis [Lumsden and Krumlauf, 1996; Capecchi, 1997; Kmita and Duboule, 2003]. In addition, *Hox* expression has also been associated with distinct pathologies like cancers [Cillo et al., 2001]. While the roles they play have been and are still under detailed investigation, the mode of action of the Hox proteins remains rather poorly documented. In particular, very few functional determinants have been mapped in Hox proteins.

The extremely conserved homeodomain encoded by the homeobox was initially described as a DNA-binding domain (DB). It is folded in three α-helices preceded by a flexible N-terminal extension. The third helix and the N-terminal extension are involved in specific contacts with the DNA base pairs through the major and minor grooves of cognate-binding sites, respectively [Gehring et al., 1994; Wolberger, 1996]. Several homeodomains have later been shown to display additional properties. In Bicoid, a Hox-related protein from *Drosophila*, the homeodomain acts as an RNA-binding domain. It interacts with the 3′untranslated region of *caudal* mRNA, and, therefore, inhibits *caudal* translation initiation [Dubnau and Struhl, 1996; Rivera-Pomar et al., 1996; Niessing et al., 2002]. Homeodomains have also been described to display determinants for intra- and intercellular trafficking. It was first reported for the Hox protein Antennapedia (Antp) that the third α-helix of the homeodomain is capable to cross-biological membranes and to be internalized in an energy- and receptor-independent manner [Derossi et al., 1996]. Later, the cell-penetrating activity that

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: FRSM; Grant number: 3.4.536.06.F; Grant sponsor: FRIA; Grant number: FC 71228; Grant sponsor: Walloon region; Grant number: 516054.

*Correspondence to: Prof. René Rezsohazy, Unit of Veterinary Sciences, Life Sciences Institute, Université catholique de Louvain, 5 (box 10) Place Croix du Sud, 1348 Louvain-la-Neuve, Belgium. E-mail: rene.rezsohazy@uclouvain.be

Received 4 September 2009; Accepted 2 February 2010 • DOI 10.1002/jcb.22563 • © 2010 Wiley-Liss, Inc.
Published online 24 March 2010 in Wiley InterScience (www.interscience.wiley.com).

the third α-helix of Antp confers to the protein or to fused cargoes, has been observed for several other homeodomain proteins [Prochiantz, 2000]. For Engrailed, another Hox-like transcription factor, it was further demonstrated that its homeodomain possesses a signal sequence for nuclear export and unconventional secretion [Maizel et al., 1999]. Finally, the homeodomain of Hox proteins has also been reported to take part to protein–protein interactions with High Mobility Group (HMG) protein 1, CREB-binding protein (CBP)/p300, Smad, Ku, Gli, or Maf proteins [e.g., Zappavigna et al., 1996; Saleh et al., 2000; Kataoka et al., 2001; Schild-Poulter et al., 2001; Shen et al., 2001; Chen et al., 2004; Shen et al., 2004; Williams et al., 2005; Li et al., 2006], or with other Hox proteins for the formation of heterodimers [Zappavigna et al., 1994].

N-terminal to the homeodomain, Hox proteins of paralog groups 1–8 contain a conserved hexapeptide sequence which is involved in the interaction with Pbx proteins and thereby plays a role in the modulation of their DNA-binding and transcriptional activity [Mann and Affolter, 1998; Moens and Selleri, 2006]. Pbx proteins are transcription factors belonging to the ''Three Amino acid Loop Extension'' (TALE) class of homeodomain transcription factors. In the hexapeptide, core tryptophan and methionine residues are particularly important to establish the Hox–Pbx contacts [Phelan et al., 1995; Passner et al., 1999; Piper et al., 1999; LaRonde-LeBlanc and Wolberger, 2003]. Accordingly, their substitution to alanine abolishes interaction between the partners and alters the specificity for DNA recognition [Phelan et al., 1995; Chang et al., 1996; Phelan and Featherstone, 1997]. In the Antp protein, the YPWM motif of the conserved hexapeptide has recently been shown to define an interface to interact with the basal transcription machinery and to stimulate target gene expression [Prince et al., 2008]. Finally, Meis and Prep, which also belong to the TALE class of homeodomain proteins, are other cofactors known to regulate Hox activity [Mann and Affolter, 1998; Jacobs et al., 1999; Shanmugam et al., 1999; Ferretti et al., 2000]. They can either directly interact with some Hox proteins in a hexapeptide independent manner or dimerize with Pbx to allow its nuclear import, therefore influencing Hox activity indirectly, through Pbx availability [Berthelsen et al., 1999; Shanmugam et al., 1999].

Functional dissections in mammalian cell transfection assays using physiological target sites or consensus-binding sites to drive reporter expression have revealed transcriptional activation domains for several Hox proteins. In the case of Hoxa5, -B1, and -D9, a transcriptional activation domain has been localized in the N-terminal region of the protein which displays some sequence conservation among Hox proteins [Zhao et al., 1996; Di Rocco et al., 1997; Vigano et al., 1998]. For HOXB3 and HOXB7, both the N- and C-termini of the protein were identified to be involved in activating transcription [Vigano et al., 1998; Chariot et al., 1999a]. More precisely, Chariot et al. [1999a] provided evidence that the N-terminal domain of HOXB7 physically interacts with CBP which in turn regulates the transcriptional properties of the Hox protein by acetylation/deacetylation of its N-terminal extremity.

Hoxa1 is one of the earliest *Hox* genes to be expressed during mouse embryonic development. It is involved in hindbrain segmentation and patterning to specify the identity of rhombomeres (r)4 and r5 as well as of neurogenic neural crest cells. Its activity is also required for normal basioccipital bone formation and for inner and middle ear development [Lufkin et al., 1991; Chisaka et al., 1992]. Recently, *Hoxa1* has also been shown to be expressed in mammary carcinoma where it appears to play a crucial role in oncogenic transformation [Chariot and Castronovo, 1996; Zhang et al., 2003].

As for other Hox proteins, little is known about Hoxa1 mode of action. Previous studies have shown that the hexapeptide of Hoxa1 is critical for its activity. In vitro, a WM-to-AA substitution in the hexapeptide prevented Hoxa1 to interact with Pbx [Phelan et al., 1995; Phelan and Featherstone, 1997] and resulted in complete loss-of-activity on distinct target enhancers [Remacle et al., 2002]. In vivo, recombinant mice harboring the WM-to-AA substitution displayed developmental defects which phenocopied those observed for the *Hoxa1* knock-out mice [Remacle et al., 2004]. This strongly suggested that the function of Hoxa1 critically relies on its interaction with Pbx. Although Hox–Pbx–Prep trimers have been observed with Hoxb1 [Di Rocco et al., 1997; Jacobs et al., 1999; Ferretti et al., 2005], no such complexes have been reported so far for Hoxa1. In addition to Pbx, Sox2, and Oct1 proteins are also known to modulate the activity of Hoxa1. These proteins affect the ability of HOXA1 and HOXB1 to stimulate *Hoxb1* expression via its autoregulatory enhancer (*b1-ARE*) during hindbrain development and, in turn, to specify the fate of the r4 territory [Di Rocco et al., 2001]. Di Rocco et al. [2001] have shown that the activity of HOXA1/Pbx complexes requires Sox/Oct binding to mediate the activation of *b1-ARE* upon retinoic acid induction, whereas HOXB1/Pbx could stimulate the *b1-ARE* in the absence of the Sox/Oct-binding site. The amino acid residues responsible for this difference of activity between HOXA1 and HOXB1 were not identified but were mapped in the N-terminal region of the proteins.

In order to identify new functional motifs important for the activity of the murine Hoxa1, we generated a set of 18 mutant Hoxa1 proteins using a random pentapeptide insertion mutagenesis approach [Hayes and Hallet, 2000]. This method has been successfully used to build up the functional map of several proteins and to generate enzyme variants with novel substrate specificities [Cao et al., 1997; Hayes and Hallet, 2000].

The mutants were assayed in different mammalian cell types for their transcriptional activity on two physiological target enhancers derived from the *EphA2* and *Hoxb1* genes. Among the 18 mutants, 11 displayed wild-type activity. For the 7 remaining mutants that showed a gain or loss-of-function, we verified their expression and intracellular localization and we further investigated their capacity to bind DNA in vitro and to activate transcription in yeast. Insertions leading to a systematic loss-of-activity either mapped to the homeodomain or to the N-terminal extremity of the protein. As expected, mutations in the homeodomain resulted in a loss of DNA-binding activity. However, the mutation in the N-terminus increased the ability of Hoxa1 to cooperatively bind DNA with Pbx. This suggests that the N-terminal extremity of the protein exerts a modulatory effect on its DNA-binding. Finally, the homeodomain and N-terminus of the protein were found to be required together to confer to Hoxa1 its intrinsic transcriptional activation potential in the yeast assay. This activation ability is otherwise modulated by the serine–threonine–proline rich C-terminal extremity of Hoxa1.

MATERIALS AND METHODS

PENTAPEPTIDE INSERTION MUTAGENESIS

We used a selectable Tn*4430* transposon (Tn*4430*::neo^R) to insert five codon long sequences into the murine *Hoxa1* open reading frame, according to the procedure described by Hallet et al. [1997]. Briefly, the Tn*4430*::neo^R and *Hoxa1* genes are harbored on separate plasmids and maintained together in *Escherichia coli* to allow jumping of Tn*4430*::neo^R into the *Hoxa1* sequence. After Tn*4430* insertion and subsequent excision by *Kpn*I digestion, 15 bp are left in the target sequence. The exact position of each Tn*4430*::neo^R insertion was determined by sequencing.

PLASMID CONSTRUCTIONS

Expression vectors for wild-type *Hoxa1* (Hoxa1^{WT}, pGIH-309), *Hoxa1*$^{QN\text{-}AA}$ (pGIH-512), and *Pbx1a* (pCMV-Pbx1a) have been previously described by Remacle et al. [2002]. Expression vector for *Prep1* was generously provided by Goudet et al. [1999]. The pAdML-ARE plasmid contains the TATA box and transcriptional start site from the Adenovirus-2 Major Late promoter (AdML), downstream of the *Hoxb1* ARE enhancer (b1-ARE [Di Rocco et al., 1997]). The EphA2-r4-Luciferase reporter plasmid (EphA2-r4) was a kind gift of Dr. Jin Chen [Chen and Ruley, 1998]. The pCMV-lacZ plasmid was previously described [Remacle et al., 2002].

The insertion mutagenesis procedure used relies on Tn*4430*::neo^R transposition into the *Hoxa1* gene, followed by *Kpn*I-mediated excision of all but 10 bp of the transposed insert. In that purpose, the unique *Kpn*I site residing in the *Hoxa1* sequence has been removed and replaced by a silent mutation generated in the *Hoxa1* expression vector pGIH-309 (*GGGTACCCC* sequence changed into *GGCT-ACCCC*; details available upon request) to obtain the pGIH900 plasmid. pGIH900 was used as a target plasmid for mutagenesis. The 15 bp insertions resulting from the transposition–excision procedure were numbered (P1–P18) according to their 5′- to3′- location in the *Hoxa1* sequence. Expression vectors for the mutant *Hoxa1* genes were named accordingly (pGIH900-1 to pGIH900-18).

For bacterial expression and protein purification, we amplified wild-type and mutated *Hoxa1* coding sequences from the pGIH-309 vector and pGIH900 series and cloned PCR fragment into the pBAD-Myc-HisA plasmid (Invitrogen) as *Nco*I–*Eco*RI fragments (primers: CCACATGTTAAACTCCTTTCTGGAATACCCCATCC and CGGAATT-CAAGTGGGAGGTAGTCAGAGTGTCTGAGG; pGIH-904 series).

For the yeast-1-hybrid assay, *Hoxa1*^{WT} and mutated *Hoxa1* coding sequences were amplified (primers: CGGAATTCTTGAA-CTCCTTTCTGGAATACCCCATCC and AAGGAAGATAAGCTAAG-AATGTGC; pGIH-905 series, details upon request) and cloned as *Eco*RI–*Pst*I fragments into the yeast vector pGBT-9 (Clontech) ensuring that the Gal-4 DB domain is fused N-terminally to the *Hoxa1* sequence.

CELL CULTURE AND TRANSIENT TRANSFECTION

ECP19 and COS7 cells were maintained in Dulbecco's modified Eagle's medium (Gibco/Invitrogen) supplemented with 7.5% (ECP19) or 10% (COS7) fetal bovine serum (Cambrex), 100 IU/ml penicillin, and 100 μg/ml streptomycin (Gibco/Invitrogen). MCF10A cells were cultured in DMEM-F12 added with 5% horse serum (Gibco/Invitrogen), 100 ng/ml cholera toxin (Tebu-Bio), 20 ng/ml hEGF, 10 μg/ml insulin, 500 ng/ml hydrocortisone (Sigma), 100 IU/ml penicillin, and 100 μg/ml streptomycin (Gibco/Invitrogen). The three cell lines were maintained at 37°C in a humidified, 5% CO_2 atmosphere.

For transient transfection, 3.6×10^5 (MCF10A), 8×10^4 (ECP19), or 1.4×10^5 (COS7) cells were plated into six-well plates. Twenty-four hours after plating, ECP19 and COS7 cells were transfected using $CaCl_2$ method [Remacle et al., 2002] and MCF10A cells were transfected with Fugene6 (Roche). Briefly, $CaCl_2$ transfections were carried out with a total amount of 10.5 μg of DNA: 3 μg of reporter plasmid, 1.5 μg of Hoxa1 expression vector, 3 μg of each Pbx1a and Prep1 expression vectors, and 0.2 or 0.02 μg of internal standard plasmid (pCMV-lacZ) in ECP19 or COS7 cells, respectively. For MCF10A transfections, a total of 2 μg of DNA (1.25 μg of reporter plasmid, 0.25μg of each expression vectors and 0.02 μg of standard plasmid) was added to 3 μl of transfection reagent.

Forty-eight hours after transfection, cells were lysed for enzymatic assays using the Luciferase Reporter Gene Assay (High Sensitivity) kit and the β-gal Reporter Gene Assay (Chemiluminescent) kit (Roche). Luciferase and β-galactosidase activities were measured with a luminometer (Glomax; Promega) and the luciferase activity was reported to the β-galactosidase activity.

WESTERN BLOTTING

COS7 cells were transfected with pGIH900 plasmid series to express Hoxa1 proteins. Forty-eight hours after transfection, cells were lysed and cytoplasmic extracts eliminated with buffer A (10 mM HEPES, 10 mM KCl, 0.1 mM EDTA, 0.05% NP40, 1 mM DTT, and 0.5 mM PMSF), and nuclear extracts were isolated in buffer B (20 mM HEPES, 400 mM NaCl, 0.1 mM EDTA, 10% glycerol, 1 mM DTT, and 0.5 mM PMSF). Nuclear fractions were run on a SDS–PAGE, blotted on nitrocellulose membrane and revealed with an anti-Hoxa1 antibody (Sigma, HPA004933). Transfection efficiency was controlled by cotransfecting a CMV-LacZ reporter and assaying the β-galactosidase activity. The protein load for Western blotting was controlled by detecting the β-actin protein (anti β-actin HRP conjugated; Sigma, A3854).

IMMUNOSTAINING AND FLUORESCENCE MICROSCOPY

MCF10A cells were seeded on glass cover slips in 24-well plates and transiently transfected with pGIH900 plasmid series to express Hoxa1 proteins. Twenty-four hours after transfection, cells were fixed in 4% formalin and blocked in 10% powder milk. Cells were incubated with the anti-Hoxa1 antibody (1/50; Sigma, HPA004933) overnight at 4°C. They were washed and incubated with a fluorescein coupled anti-rabbit IgG antibody (1/100; GE Healthcare N1034) for 1 h. Cover slips were mounted in vectatshield with Dapi medium (Vector Laboratories H1200) and viewed under a Polyvar microscope (Reichert Jung).

PROTEIN PURIFICATION

Myc-His-tagged Hoxa1 proteins were expressed in *E. coli* TOP10 bacteria using the pBAD-Myc-HisA vector. Protein expression was

induced by adding 2% L-(+)-arabinose (Sigma) during the exponential phase of growth (at $A_{600\,nm} \approx 0.5$). After 6 h of induction, bacteria were harvested, lysed, and protein purification was performed using TALON columns (Clontech) according to the manufacturer instructions. Elutions were performed with 150 mM imidazole.

ELECTROPHORETIC MOBILITY SHIFT ASSAY

Hoxa1 proteins were purified as described above. Pbx1a and Prep1 proteins were produced using the in vitro transcription/translation TnT coupled reticulocyte lysate system (Promega) as described by the manufacturer. Oligonucleotides used were derived from the *Hoxb1 ARE Repeat 3* sequence (PM1-R3) which comprises Hox–Pbx and Prep-binding sites (TCTTTGTCATGCTAATGATTGGGGGGT-GATGGATGGGCGCTG; [Ferretti et al., 2005]) and from the *Repeat E* of the *EphA2-r4* enhancer (R_E; TTGCATGATG-GATGGGCTGG; [Chen and Ruley, 1998]). Oligonucleotides were labeled with γ-^{32}P-ATP using the T4 Polynucleotide Kinase (Biolabs), purified on a ChromaSpin + TE-10 column (Clontech) and hybridized with the complementary oligonucleotide by heating at 95°C for 5 min and slowly cooling at room temperature. Proteins were incubated with probes and poly dI-dC (25 μg/ml) for 30 min at room temperature and for 30 min on ice in binding buffer (10 mM Tris–HCl pH 7.5, 75 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 mg/ml BSA, and 12% glycerol). Samples were then resolved by electrophoresis on a non-denaturing 5% polyacrylamide gel in 1× TBE. Gels were dried and exposed to X-rays film at −80°C. To use equivalent amounts of the Hoxa1 protein variants in the EMSA, a semi-quantitative estimate of purified Hoxa1 proteins was performed by Western blotting and autoradiography scanning (Kodak 1D3.5 program).

YEAST 1-HYBRID ASSAY

Yeast strain PJ696 (genotype: trp1–901, leu2–3, 112, ura3–52, his3–200, gal4Δ, gal80Δ, GAL2-ADE2, LYS2::GAL1-HIS3, and met2::GAL7-lacZ) was transformed with the plasmids coding for the Gal-4 DB-Hoxa1 fusions using a lithium acetate method [Chen et al., 1992] and selected on synthetic dextrose media (SD: 2% glucose, 0.7% yeast extract w/o aa) supplemented with methionine (M), uracile (U), lysine (K), histidine (H), adenine (A), and leucine (L). The transcriptional activity of Gal-4 DB-Hoxa1 was tested following the expression of reporter genes *His3*, *Ade2*, and *LacZ*. Expression of the first two markers was examined by a complementation assay on media lacking histidine (SD + MUK + LA), adenine (SD + MUK + LH) or both (SD + MUK + L). For the enzymatic assay, yeast cells were lysed (Tris–HCl 100 mM, DTT 1 mM, glycerol 20%, glass beads) and β-galactosidase activity was determined with the β-gal Reporter Gene Assay (Chemiluminescent) kit (Roche). β-galactosidase activity was normalized to total proteins, as dosed by Bradford protein assay (Sigma).

STATISTICS

Statistical analysis of enzymatic activity data was performed according to a Student's *t*-test (GraphPadPrism 4).

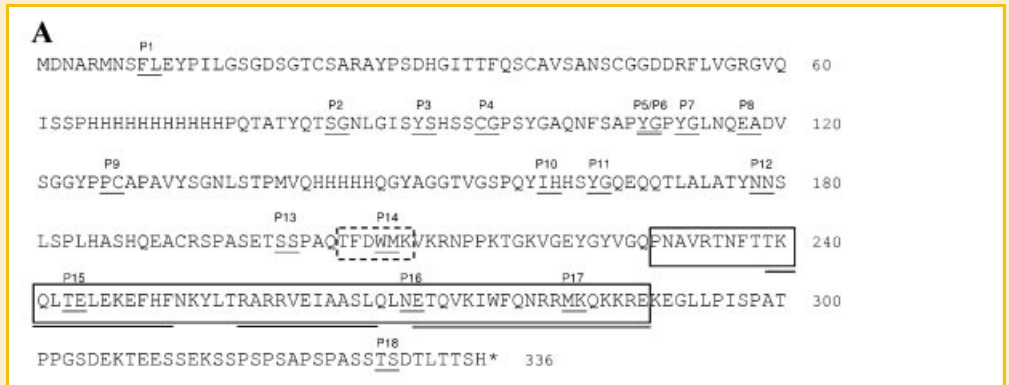
RESULTS

Hoxa1 IS PERMISSIVE TO PENTAPEPTIDE INSERTIONS OUTSIDE THE HOMEODOMAIN AND N-TERMINAL EXTREMITY

To identify new functional determinants in the Hoxa1 transcription factor, we generated a series of mutants by a random pentapeptide insertion method. Briefly, this method relies on the insertion of the bacterial transposon Tn*4430* followed by its incomplete excision by restriction cleavage. Tn*4430* contains *Kpn*I cleavage sites 5 bp inwards from both its termini. In addition, Tn*4430* duplicates 5 bp of the target sequence upon transposition. Therefore, deletion of the transposon from the target gene with *Kpn*I leaves an in-frame insertion of 15 bp composed of 10 bp from Tn*4430* and 5 bp of duplicated target DNA. This in turn results in 5 amino acids insertion in the target protein. Due to Tn*4430* sequence, the inserted pentapeptides contain at least one proline residue known to disorganize secondary structures of proteins [Hallet et al., 1997] (Fig. 1).

We generated 18 insertion mutant derivatives of Hoxa1. Insertions were randomly spread along the protein sequence (Fig. 1). Among the 18 mutations, four hit well characterized domains of the protein. Three insertions were localized in the homeodomain (mutants P15, P16, and P17) and one in the hexapeptide motif involved in Pbx binding (mutant P14). One insertion fell at the N-terminus of the protein (mutant P1), in a region that shares sequence conservation between Hoxa1 orthologs in vertebrates and with other Hox proteins of the paralogy group 1 (Supplementary Fig. S1 and S2). Another insertion was located in a serine–threonine–proline rich region conserved at the C-terminus of known vertebrate Hoxa1 proteins (mutant P18, Supplementary Fig. S1). Two additional insertions occurred between residues that are highly conserved among Hoxa1 sequences, but in regions of unknown function with no particular amino acid sequence composition (mutants P4 and P11). The remaining insertions were found between less conserved residues in different parts of the protein.

To evaluate the effect of pentapeptide insertions on the transcriptional activity of Hoxa1, we tested all the mutants in co-transfection assays. We used three different cell lines to address the influence of the cell context on the activity changes displayed by the mutants: MCF10A, a human mammary epithelial cell line; COS7, a Rhesus monkey kidney cell line; and ECP19, a mouse embryonic carcinoma cell line. The activity of Hoxa1 proteins was evaluated by their ability to activate a luciferase reporter gene controlled by the "r4" enhancer of the *EphA2* gene (*EphA2-r4*). *EphA2* is expressed in the developing hindbrain and its r4-specific expression is reminiscent of that of *Hoxa1* and *Hoxb1*. *EphA2* expression has moreover been shown to be decreased in *Hoxa1*/*Hoxb1* mutant mice [Chen and Ruley, 1998], suggesting that it is a physiological target of Hoxa1 and/or Hoxb1. The *EphA2-r4* enhancer was further found as a Hoxa1 responsive regulatory module that contains five Hox- and Pbx-binding sites [Chen and Ruley, 1998] and at least one Meis/Prep-binding site allowing synergistic association of Meis/Prep with Pbx and Hox. Expression vector for each Hoxa1 mutant was co-transfected with the *EphA2-r4-luciferase* reporter plasmid. As expected, while the Pbx1a and Prep1 proteins did not provide



Mutants	Nucleotidic insertions	Amino Acid insertions
P1	*TTT CT*G GGG TAC CCC TTT CT	Phe_{9} Leu Gly Tyr Pro Phe
P2	*T TCT C*GG GGT ACC CCT TCT G	Ser_{84} Gly Gly Thr Pro Ser
P3	*T TAT T*GG GGT ACC CCT TAT T	Tyr_{91} Trp Gly Thr Pro Tyr
P4	*T TGT G*GG GGT ACC CCT TGT G	Cys_{96} Gly Gly Thr Pro Cys
P5	*CT TAT* GGG GTA CCC CCT TAT	Tyr_{108} Gly Val Pro Pro Tyr
P6	*T TAT G*GG GGT ACC CCT TAT G	Tyr_{108} Gly Gly Thr Pro Tyr
P7	*CC TAT* GGG GTA CCC CCC TAT	Tyr_{112} Gly Val Pro Pro Tyr
P8	*AG GAA* GGG GTA CCC CAG GAA	Glu_{117} Gly Val Pro Gln Glu
P9	*CC CCG* GGG GTA CCC CCC CCG	Pro_{126} Gly Val Pro Pro Pro
P10	*ATT CA*G GGG TAC CCC ATT CA	Ile $_{161}$ Gln Gly Tyr Pro Ile
P11	*A TAT G*GG GGT ACC CCA TAT G	Tyr_{165} Gly Gly Thr Pro Tyr
P12	*T AAT A*GG GGT ACC CCT AAT A	Asn_{178} Arg Gly Thr Pro Asn
P13	*T TCT C*GG GGT ACC CCT TCT C	Ser_{201} Arg Gly Thr Pro Ser
P14	*TGG AT*G GGG TAC CCC TGG AT	Trp_{207} Met Gly Tyr Pro Trp
P15	*TC ACA* GGG GTA CCC CTC ACA	Thr_{242} Gly Val Pro Leu Thr
P16	*TC AAT* GGG GTA CCC CTC AAT	Asn_{268} Gly Val Pro Leu Asn
P17	*ATG AA*G GGG TAC CCC ATG AA	Met_{281} Lys Gly Tyr Pro Met
P18	*T ACC T*GG GGT ACC CCT ACC T	Thr_{326} Trp Gly Thr Pro Thr

Fig. 1. Hoxa1 insertion mutants. A: The primary sequence of the Hoxa1 protein is shown and amino acid residues between which the pentapeptide insertions are located are underlined. Numbering of the corresponding mutants (P1–P18) is indicated above the sequence. The homeodomain sequence is boxed, with the sequence of the α-helices underlined. A dashed box indicates the hexapeptide. B: The oligonucleotide and corresponding pentapeptide insertions are presented for each Hoxa1 mutant. Nucleotide insertions (bold case) correspond to a common core of 10 bp from the Tn*4430* transposon (GGGGTACCCC) and to five nucleotides duplicated from the *Hoxa1* sequence. Nucleotides are presented in triplets according to the *Hoxa1* open reading frame. The coordinate of the pentapeptide insertion (bold case) is indicated for each mutant according to the amino acid located N-terminally to the insertion.

reporter activation alone (Fig. 2A,B), $Hoxa1^{WT}$ was able to activate reporter activity in the presence of Pbx1a in MCF10A cells and COS7 cells. This activity was further enhanced by addition of Prep1 (Fig. 2A,B). In the ECP19 cells, Hoxa1 did not require co-transfection of Pbx1a and Prep1 to activate the reporter (Fig. 2C) because these cofactors were endogenously expressed at a sufficient level to provide full Hoxa1 activity (data not shown). Indeed, co-transfection of Pbx1a and Prep1 expression vectors did not enhance Hoxa1-mediated reporter activation in this cell line.

Co-transfection experiments in the three cell lines showed that the pentapeptide insertions differently affected the activity of Hoxa1 (Fig. 3). The mutant targeted at the N-terminal extremity (mutant P1) and the three mutants in the homeodomain (mutants P15, P16, and P17) showed a significant loss-of-function in the three cell lines (Fig. 3) indicating that the corresponding insertions impaired protein activity. Residual transcriptional activity detected for these four mutant proteins was in the same range as that measured for the DNA-binding null mutant $Hoxa1^{QN\text{-}AA}$ in which the homeodomain residues Gln50 and Asn51 have been substituted to Ala [Matis et al., 2001; Remacle et al., 2002]. Insertion near the hexapeptide (mutant P13) resulted in lower activity in MCF10A and COS7 cells (Fig. 3A,B) but in a gain-of-activity in ECP19 cells (Fig. 3C). Similarly, mutant P5 with an insertion at the middle of the protein showed a loss-of-function in MCF10A cells (Fig. 3A) and a gain-of-function in ECP19 cells (Fig. 3C). Three additional mutants (mutants P2, P6, and P18) displayed stronger activity in ECP19 cells (Fig. 3C) but behaved like $Hoxa1^{WT}$ in the other cell types. Mutant P3 activity was decreased with respect to the wild-type protein, but this was only observed in COS7 cells (Fig. 3B). Finally, eight mutants (mutants P4, P7, P8, P9, P10, P11, P12, and P14) displayed similar transcriptional activity as $Hoxa1^{WT}$ in the three cell lines (Fig. 3).

A striking observation is that five mutants displayed a gain-of-function and that this was only observed in ECP19 cells. To confirm this, the 10 mutants showing a gain or loss-of-activity on *EphA2-r4* enhancer in any of the three cell models were assayed again in ECP19 cells but on another target enhancer (Fig. 3D). The *b1-ARE*

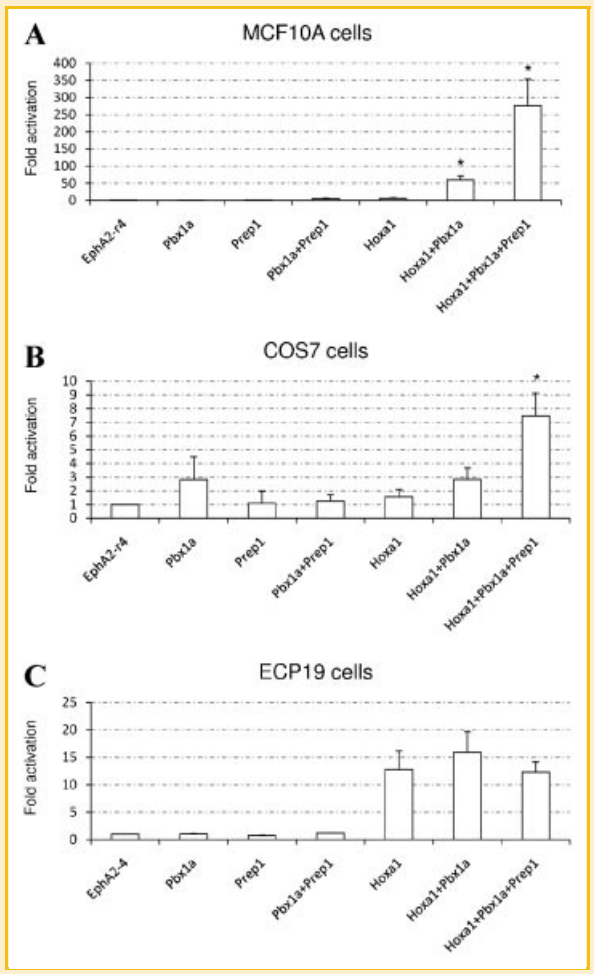


Fig. 2. Transfection assays for the Hoxa1 activity on the *EphA2-r4* enhancer. MCF10A cells (A), COS7 cells (B), and ECP19 cells (C) have been transfected for the *EphA2-r4-luciferase* reporter construct, alone (EphA2-r4) or together with expression vectors for Hoxa1, Pbx1a, and/or Prep1 proteins. Values are expressed as fold reporter activation, the standardized luciferase activity provided by the *EphA2-r4-luciferase* reporter alone corresponding to 1. Bars indicate the standard deviation of 5 (A,B) or 2 (C) independent experiments. $^{*}t < t_{0.05}$.

enhancer isolated from the *Hoxb1* gene is a physiological target of Hoxa1 involved in the cross-talk between *Hoxa1* and *Hoxb1* in the developing hindbrain [Studer et al., 1998]. It contains three binding sites for Pbx–Hox heterodimers and a Meis/Prep-binding site [Ferretti et al., 2005]. The mutant affected in the N-terminus region of the protein (mutant P1) and two mutants in the homeodomain (mutants P15 and P17) presented a loss-of-function. Although not statistically validated ($P = 0.06$), activity of the third homeodomain mutant (mutant P16) was also affected. Transfections performed with the *b1-ARE* enhancer confirmed the gain-of-function observed in ECP19 cells for three mutants (mutants P5, P13, and P18), but not for mutants P2 and P6. Finally, mutant P3 showed wild-type activity on *b1-ARE*.

Taking these results together, we selected seven mutants displaying altered activity on both *EphA2-r4* and the *b1-ARE* enhancers for further analysis: mutants P1, P15, P16, and P17 which showed a loss-of-function under any condition; mutant P18 which presented a gain-of-function on both enhancers in ECP19 cells and mutants P5 and P13 which showed a gain-of-function for both enhancers in ECP19 cells but a loss-of-function on *EphA2-r4* in COS7 (P13) or MCF10A (P5 and P13) cells.

All the selected mutants displayed a low, background activity in the absence of Pbx1a and Prep1 in MCF10A and COS7 cells (Supplementary Fig. S3), their loss-of-function being conspicuous only in the presence of the cofactors. In the ECP19 line, co-transfection of Pbx1a and Prep1 expression vectors did not rescue the loss-of-function phenotype provided by mutants P1, P15, P16, and P17, suggesting these are constitutive loss-of-function Hoxa1 variants. However, by overexpressing Pbx1a and Prep1, the gain-of-function observed in ECP19 cells for the P5, P13, and P18 mutants was suppressed, confirming that their gain-of-activity is conditional and context dependent (Supplementary Fig. S3).

Western blot detection of the Hoxa1 variants showed two specific bands that are typical of Hoxa1 [Zhang et al., 2003] indicating that all the mutants were properly and stably expressed in co-transfection assays (Fig. 4). Small variations in the relative signal intensity of the two bands did not correlate with the different levels of activity displayed by the mutants. Because an intracellular relocation of the proteins could explain the observed changes in activity, we documented the intracellular distribution of the Hoxa1 variants by immunocytochemical detection. As expected from its role as a transcription factor, Hoxa1 was predominantly detected in the nucleus of transfected MCF10A cells. All the mutant proteins also localized in the nucleus indicating that the different activity displayed by Hoxa1^{WT} and the mutants was not a consequence of an altered subcellular localization (Supplementary Fig. S4).

All together these data show that the Hoxa1 protein is highly permissive to mutagenesis. Indeed, only four insertions were detrimental for Hoxa1 activity, and strikingly, three of them fell into the only known globular domain of the protein, the homeodomain. The other insertions either did not affect the activity of Hoxa1 at all or led to moderate modulation of its function depending on the enhancer or the cell line used in transfection experiments.

THE TRANSCRIPTIONAL ACTIVATION CAPACITY OF Hoxa1 RELIES ON THE HOMEODOMAIN AND THE N-TERMINAL REGION AND IS MODULATED BY THE SERINE–THREONINE–PROLINE RICH C-TERMINAL EXTREMITY

Yeast is a suitable model to evaluate the transcriptional activation capacity of a transcription factor. By fusing Hoxa1 to the Gal-4 DB, the hybrid protein binds DNA at UAS elements recognized by its Gal-4 moiety. In that context, downstream reporter genes would be activated if Hoxa1 possesses the intrinsic ability to contact and stimulate the transcription machinery. The PJ696 yeast strain contains three reporter genes placed under the control of UAS elements for Gal-4: *His3*, *Ade2*, and *LacZ*. When expressed, the first two genes allow yeast to grow on selective media devoid of histidine and/or adenine. The third reporter provides a mean to quantify transcriptional transactivation by measuring the induced β-galactosidase activity.

The growth assay revealed that Hoxa1^{WT} is able to stimulate reporters in yeast meaning that the protein is able to activate the

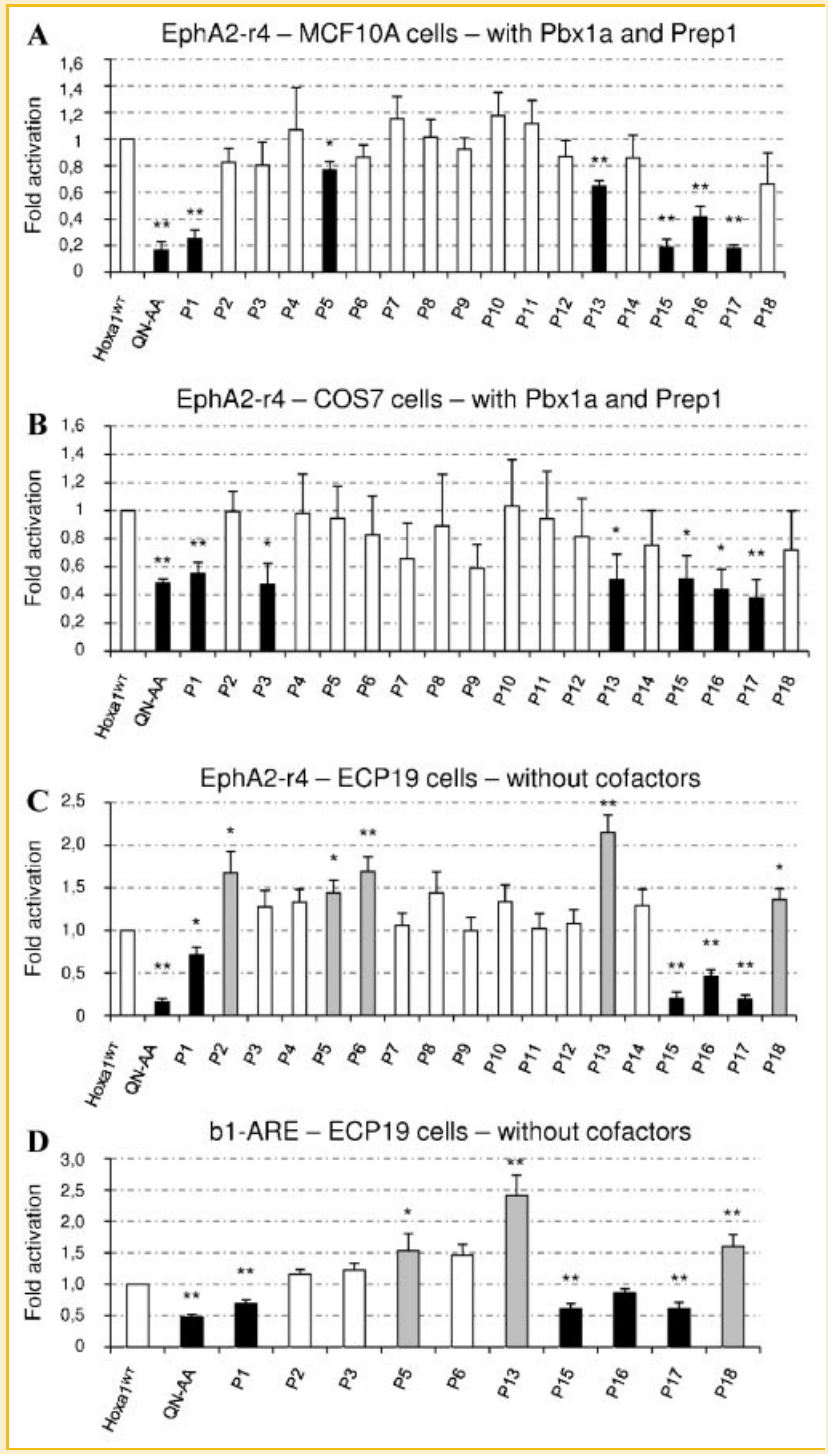


Fig. 3. Comparison between the wild-type and mutant Hoxa1 activity. MCF10A cells (A), COS7 cells (B), and ECP19 cells (C,D) have been transfected with either the *EphA2-r4-luciferase* reporter plasmid (EphA2-r4, A–C) or the *b1-ARE* reporter (b1-ARE, D) together with the expression vector for Hoxa1 ($Hoxa1^{WT}$), the QN-to-AA homeodomain Hoxa1 mutant (QN-AA), or the pentapeptide insertion mutants (P1–P18). Expression vectors for Pbx1a and Prep1 were co-transfected in MCF10A and COS7 cells while not in ECP19 cells. Values are expressed as fold reporter activation with respect to that provided by the $Hoxa1^{WT}$ which is taken as reference (value of 1). Black boxes indicate activities that are significantly lower than that of the wild-type and gray boxes indicate significantly higher activities. Bars indicate the standard deviation of 4 (A), 6 (B), 8 (C), or 12 (D) independent experiments. $^{*}t < t_{0.05}$ and $^{**}t < t_{0.01}$.

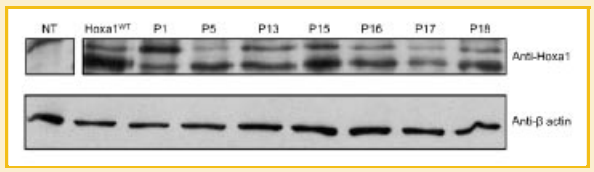


Fig. 4. Western blot detection of the Hoxa1 protein variants. Nuclear extracts of MCF10A cells transfected for an expression vector for Hoxa1 or its insertion derivatives (P1, P5, P13, P15, P16, P17, P18) were loaded on gel and transfered for Western blot analysis with anti-Hoxa1 antibody. Identical sample loads were processed for detection of the constitutively expressed β-actin protein. Two specific bands for Hoxa1 are revealed in each sample except for the non-transfected sample (NT).

basic transcription machinery. However, the mutant affected in the N-terminal region (mutant P1) and the three mutants with insertions in the homeodomain (mutants P15, P16, and P17) failed to complement growth on selective media. Conversely, mutant P18 grew faster on selective media (data not shown). These data were confirmed by measuring the β-galactosidase reporter activity (Fig. 5A). For the first four mutants (P1, P15, P16, and P17) a dramatic decrease in β-galactosidase activity was observed while the C-terminal P18 mutant stimulated *LacZ* at a 30% higher level than Hoxa1^{WT} (Fig. 5A).

Deletion derivatives of Hoxa1 were then generated for which 46 aa and 47 aa residues were removed from the N- and C-terminus of the protein, respectively. The N-terminal deletion resulted in a

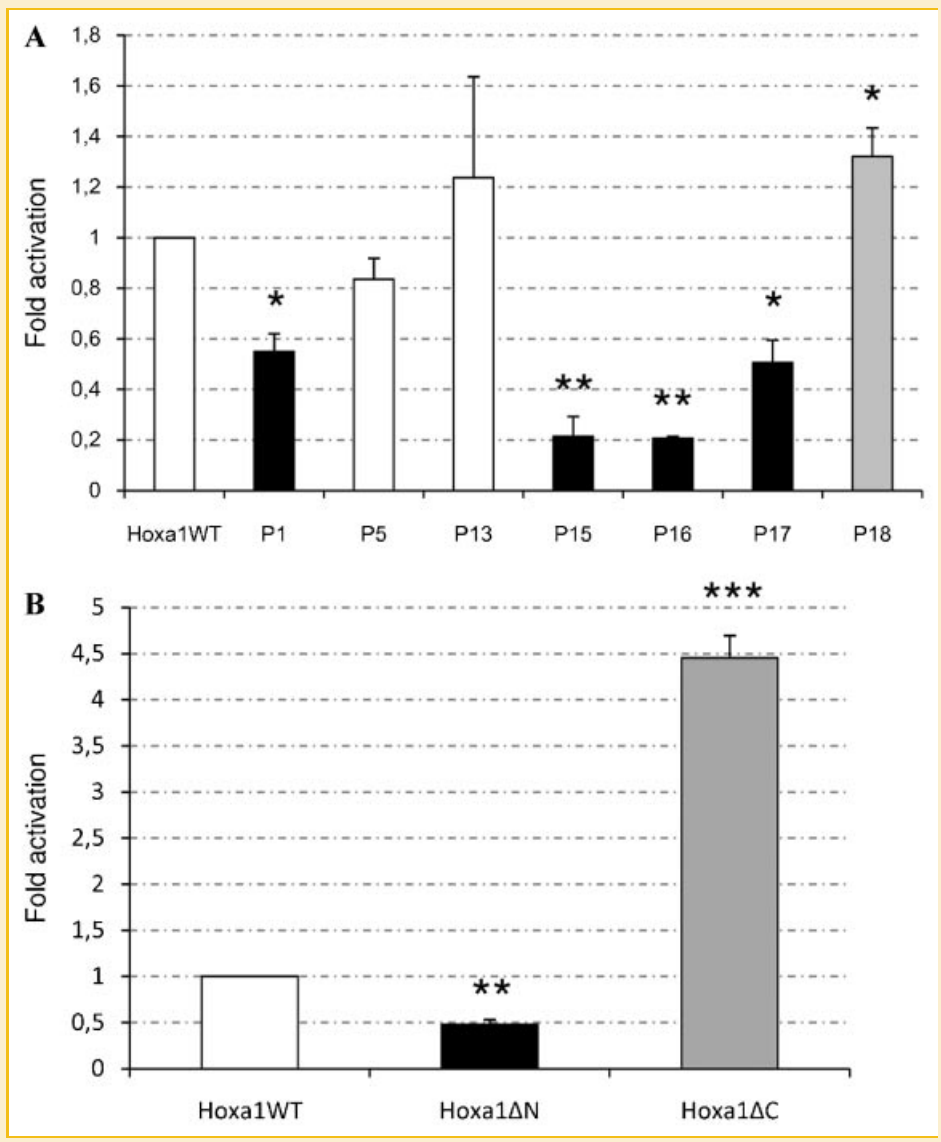


Fig. 5. Transcriptional activity of Hoxa1 variants in yeast one-hybrid assay. Yeast cells were transformed with expression vectors for Gal4-DB-Hoxa1 fusion proteins and assayed for the activity of a *LacZ* reporter gene. β-galactosidase activities were normalized by quantification of the total proteins in the extracts, and the activation provided by the wild-type Hoxa1 was referred to as 1. The activity provided by the wild-type Hoxa1 protein is compared to that measured for the pentapeptide insertion mutants P1, P5, P13, P15–18 (A) and for its N-terminal (Hoxa1ΔN) and C-terminal (Hoxa1ΔC) deletion derivatives (B). Black and gray boxes indicate significant loss or gain-of-activity as compared to Hoxa1^{WT}, respectively. Bars indicate the standard deviation of four independent experiments. $^{*}t < t_{0.05}$; $^{**}t < t_{0.01}$; $^{***}t < t_{0.001}$.

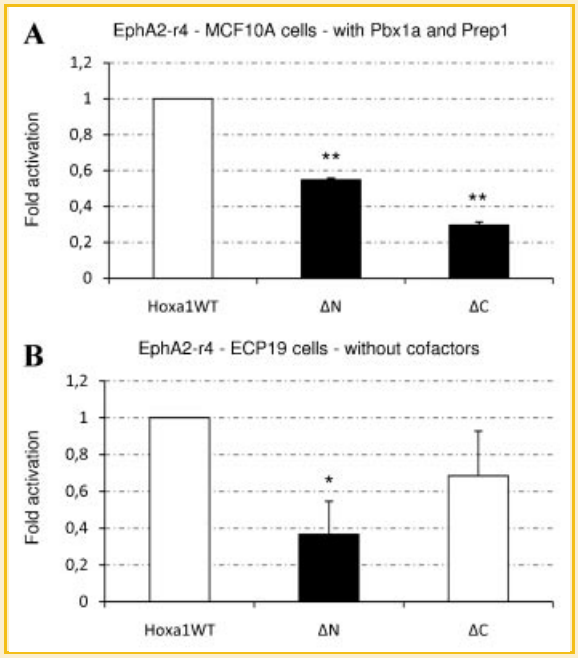


Fig. 6. Comparison between the wild-type Hoxa1 and deletion derivatives activity. MCF10A cells (A) and ECP19 cells (B) have been transfected with the *EphA2-r4-luciferase* reporter plasmid together with the expression vector for Hoxa1 (Hoxa1^{WT}) or the N-terminally (ΔN) or C-terminally (ΔC) deleted mutants insertion. Expression vectors for Pbx1a and Prep1 were co-transfected in MCF10A cells while not in ECP19 cells. Values are expressed as fold reporter activation with respect to that provided by the Hoxa1^{WT} which is taken as reference (value of 1). Black boxes indicate activities that are significantly lower than that of the wild-type. Bars indicate the standard deviation of four independent experiments. $^{*}t < t_{0.05}$; $^{**}t < t_{0.01}$.

complete loss-of-transactivation, whereas the C-terminal deletion led to a fourfold increase in activity (Fig. 5B). These deleted Hoxa1 variants were also tested in co-transfection assays, which confirmed that the N-terminus of Hoxa1 is required for its wild-type activity. However, the C-terminal truncation resulted in a cell-type dependent phenotype confirming its conditional modulatory effect on the Hoxa1 activity, as it was observed for the P18 C-terminal insertion (Fig. 6).

Therefore, both the N-terminal region and the homeodomain of Hoxa1 are required to confer its transcription activation potential which is otherwise modulated by its serine–threonine–proline rich C-terminus.

THE Hoxa1 N-TERMINAL EXTREMITY MODULATES Hoxa1–Pbx1a COMPLEX FORMATION ON TARGET DNA

Electrophoretic mobility shift assays were performed to determine whether the selected mutants showed altered DNA-binding capacity in the presence of Pbx1a or Pbx1a and Prep1 cofactors. For these experiments, we examined binding of purified Hoxa1^{WT} and mutant proteins on two different DNA probes. The first probe corresponds to the *b1-ARE* repeat three element (PM1-R3; [Ferretti et al., 2005]) which contains Pbx–Hox and Prep recognition sites. The second probe contained the *EphA2-r4* repeat E (R_E) element that displays a single Pbx- and Hox-binding site [Chen and Ruley, 1998]. Hoxa1 alone was unable to bind either probe to a detectable level (data not shown). On the PM1-R3 probe, Pbx1a and Prep1 formed a dimer in the absence of Hoxa1 (Fig. 7A). Upon addition of Hoxa1, we observed an additional shifted band corresponding to the Hoxa1–Pbx1a–Prep1 trimeric complex (Fig. 7A). On the R_E probe, Pbx1a did not bind DNA alone but formed a co-complex with Hoxa1 (Fig. 7B).

As expected, mutants affected in the homeodomain completely lost the capacity to bind DNA on both the PM1-R3 and R_E probes (mutants P15, P16, and P17; Fig. 7A,B). Mutants P5, P13, and P18 behaved like the wild-type protein, forming the same levels of co-complexes with Pbx1a and Pbx1a–Prep1 when equivalent amounts of proteins were added to the reactions (see Materials and Methods Section; Fig. 7A,B). Thus, the corresponding pentapeptide insertions did not modify the capacity of Hoxa1 to bind DNA or to interact with Pbx1a. In sharp contrast, the Hoxa1 variant affected in its N-terminal region (mutant P1) showed an increased ability to synergize with Pbx1a or Pbx1a–Prep1 complex upon DNA-binding (Fig. 7). For the PM1-R3 probe, the equilibrium between the dimeric Pbx1a–Prep1 DNA complex and the Hoxa1–Pbx1a–Prep1 trimer complex was shifted towards formation of the trimer complex in reactions performed with the mutated protein P1 (Fig. 7A). Likewise, incubation of the R_E probe with the P1 mutant yielded higher levels of Hoxa1–Pbx1a co-complexes than observed with equivalent amounts of the wild-type Hoxa1 (Fig. 7B).

The higher stability of protein–DNA complexes containing the P1 Hoxa1 mutant was confirmed by competition experiments using unlabeled probes as a specific competitor (Fig. 7C). With the wild-type protein, addition of a 5× to 10× excess of cold PM1-R3 probe resulted in a rapid and complete dissociation of the trimeric complex, while the Pbx1a–Prep1 dimer complexes remained detectable in the binding reactions. With the P1 mutant, the trimer–DNA complex was not titrated out at a lower competitor concentration than the dimer complex (Fig. 7C). Similarly, a higher concentration of unlabeled R_E probes was needed to totally dissociate the complex formed with mutant P1 than with Hoxa1^{WT} (data not shown). Finally, like the wild-type protein, mutant P1 did not bind to either probes in the absence of cofactors (not shown) suggesting that the gain in protein–DNA complex formation displayed by mutant P1 requires its Pbx1a partner. This therefore suggests that the N-terminal extremity of Hoxa1 exerts a negative effect on Hox–Pbx DNA-binding cooperativity that is relieved by the P1 pentapeptide insertion. Alternatively, the pentapeptide insertion might facilitate a structural transition required for Hoxa1 to bind DNA together with Pbx1a.

DISCUSSION

In this study, we generated 18 pentapeptide insertion mutants to identify functional determinants of the Hoxa1 transcription factor. About two-third of the mutations had no apparent effect on Hoxa1 activity and only four insertions led to a severe loss-of-function. This shows that Hoxa1 is remarkably permissive to oligopeptide insertions. By comparison, similar studies based on the same mutagenesis strategy typically yielded ~10–20% of unaffected

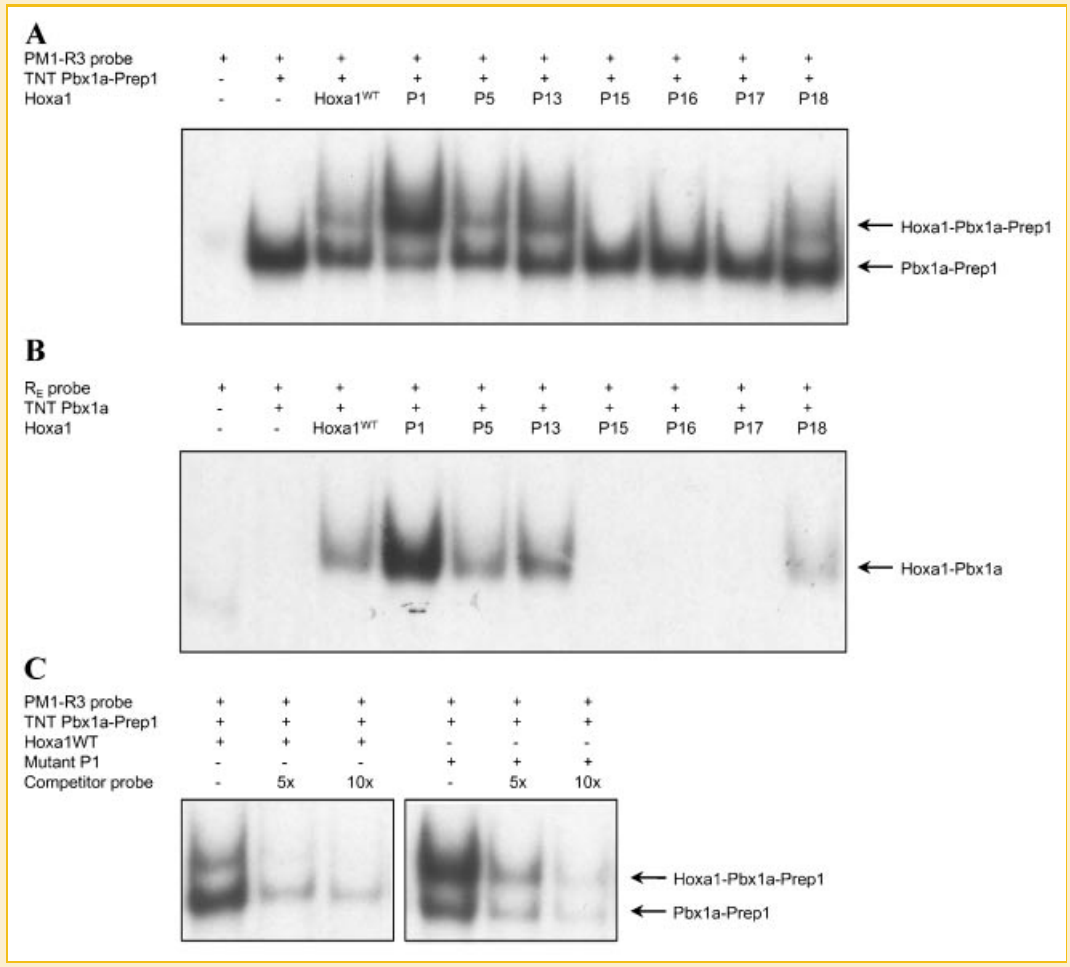


Fig. 7. DNA binding of Hoxa1^{WT} and pentapeptide insertion mutants. EMSA were run after incubating (A) the *Hoxb1 ARE* repeat 3 (PM1–R3) or (B) *EphA2–r4* repeat E (R_E) probes together with in vitro translated Pbx1a and Prep1 and purified Hoxa1 proteins (Hoxa1^{WT} and P1, P5, P13, P15, P16, P17, P18 insertion mutants). Competition assays were performed for Hoxa1 and the P1 mutant by incubating probe PM1–R3 (C) and proteins together with increasing amounts of corresponding unlabeled competitor probes. Pbx1a–Prep1 dimer and Pbx1a–Prep1–Hoxa1 trimer complexes bound to DNA are indicated (arrows).

mutants, while the vast majority of inserted oligopeptides altered the protein function to various extents depending on the location and the type of secondary structure that was affected by the insertion [Hayes and Hallet, 2000 and references therein]. As all inserted pentapeptides contained at least one proline residue, the exceptionally high permissiveness revealed here for Hoxa1 suggests that the protein exhibits little secondary structure elements or specific folds, except for the homeodomain which received three out of the four fully deleterious mutations. This is consistent with structure modeling of Hoxa1 predicting the presence of many disordered regions outside the homeodomain (data not shown). High degree of intrinsic disorder has been proposed to be shared by many transcription factors [Liu et al., 2006]. For example, beside their well structured DNA-binding domain, the activation domains of CREB and p53 switch from a disordered to an ordered conformation depending on their binding partner [Liu et al., 2006]. In that regard, transcription factors can be viewed as DNA-binding platforms involved in various contextual interactions with DNA and multiple transcription factors, coregulators, etc. Disordered regions provide the required flexibility allowing transcription factors making complex network of interactions with many different targets.

Functional tests in transfected animal cells or in yeast demonstrated that the homeodomain of Hoxa1 is not only required for DNA-binding, but also for the activation of transcription. In yeast, a Gal4-DB-Hoxa1^{WT} fusion was able to stimulate reporter gene transcription showing that Hoxa1 possesses the intrinsic potential for transcriptional transactivation. However, all three mutants affected in the homeodomain were impaired for this ability. The homeodomain of other Hox or Hox-related proteins was also shown to contribute to transcriptional activity. In most cases, however, the homeodomain appeared to exert a negative influence on transcription in the context of synthetic or physiological target enhancers in transfected mammalian cells [Schnabel and Abate-Shen, 1996; Schild-Poulter et al., 2001; Shen et al., 2001].

The transactivation potential of Hoxa1 also requires the N-terminal extremity of the protein which displays sequence conservation with several other Hox proteins [Tour et al., 2005]. A pentapeptide insertion in this region (mutant P1) led to a loss-of-

activity in transfected cells and in yeast assays. Deletion of the Hoxa1 N-terminus also led to a loss of transcriptional activity in the yeast assay. Consistent with our data, it has been shown that the N-terminus of HOXB1, the closest HOXA1 homolog, is also required for transcriptional activation [Di Rocco et al., 2001]. However, the functional determinant involved in that transcription activity has not been finely mapped. The N-termini of HOXB7, HOXD4, HOXD9, HOXB3, and the *Drosophila* Ubx and Scr proteins also share the sequence conservation and have been reported to be involved in their transactivation potential [Vigano et al., 1998; Chariot et al., 1999a; Saleh et al., 2000; Tour et al., 2005]. Again, except for the *Drosophila* Hox proteins that were assayed in vivo, the conclusions for mammalian Hox were basically drawn from in vitro data but mostly from assays based on physiologically validated target sequences. For HOXB7 and HOXD4, the N-terminal domain acts by the recruitment of the general co-regulator CBP histone acetyl-transferase. Otherwise, the activation provided by the N-terminal determinants can also be mediated, at least in part, by the interaction with specific partner proteins. For example, the N-terminus of HOXB7 interacts with IkB-α which stimulates its transactivation activity [Chariot et al., 1999b]. In that context, it is worth mentioning that the N-terminal extremity of HOXA1 and HOXB1 has been pinpointed for the differential activity these proteins display in association with Sox2 and Oct-1 proteins [Di Rocco et al., 2001]. However, the loss-of-transactivation observed for the Hoxa1 mutants in yeast must depend on more general factors that are conserved in the yeast model. Altogether our data provide that the interaction with the transcription machinery or with general transcription factors involves both the homeodomain and N-terminal regions of Hoxa1. Whether these two regions are brought together to define an interaction interface or whether they independently interact with basic transcription factors needs further functional and structural investigation.

The pentapeptide insertion obtained in the serine–threonine–proline rich C-terminus of the protein enhanced the transactivation mediated by the Gal4-DB–Hoxa1 hybrid in yeast and also resulted in a gain-of-activity of the protein in co-transfection assay in ECP19 cells. Consistently, removal of 47 aa residues from the C-terminus significantly enhanced the Hoxa1 activity in the yeast assay. However, this deletion provoked a context-dependent loss-of-activity in mammalian cells. The transactivation potential of Hoxa1 appears then to be modulated by this protein motif. The serine–threonine–proline rich sequence of the Hoxa1 C-terminus is conserved among the vertebrate Hoxa1 homologs (Supplementary Fig. S1), and is predicted to be disordered and to display phosphorylation sites for CKII and PKC (data not shown). The influence that this region displays on Hoxa1 activity may therefore be altered by post-translational modification. Interestingly, some predicted phosphorylation sites in the Hoxa1 sequence are lost in the C-terminal mutant (P18). The serine–threonine–proline rich C-terminal domain of Hoxa1 is not conserved in other Hox proteins than Hoxa1 orthologs. However, serine- and/or proline-rich protein regions N-terminally located with respect to the homeodomain have been found to cause repression of transcription by HOXA7 and HOXC8 [Schnabel and Abate-Shen, 1996] and a C-terminal serine–threonine rich domain was also reported to contribute to the transcription activity of the *Drosophila* Hox protein Ubx [Ronshaugen et al., 2002].

The pentapeptide insertion at the N-terminal end of Hoxa1 increases its ability to make DNA-binding complexes with Pbx1a protein. This means that besides its involvement in transcription activation, a novel structural and functional implication is uncovered for the conserved N-terminus as it influences cooperativity between Hoxa1 and Pbx1a for DNA-binding. The fact that, on the one hand, the N-terminal mutant better synergizes with Pbx1a for their interaction on DNA, and, on the other hand, has a decreased transactivation potential might be functionally linked. It was proposed that the conformational changes induced upon Hox–Pbx interaction facilitate an activity switch of the complex so as to stimulate transcriptional activation [Pinsonneault et al., 1997; Merabet et al., 2003]. Such a structural and functional link between modulation of DNA-binding and transcription activation has already been highlighted for an intrinsically disordered region of the Ubx protein [Liu et al., 2008]. In Hoxa1, while the N-terminal insertion may facilitate the conformational transition required for Hox–Pbx DNA-binding, it may at the same time impair its subsequent interaction with the transcriptional machinery.

Interestingly, one pentapeptide insertion was obtained in the hexapeptide motif involved in the interaction with Pbx. This insertion did not affect the activity of the protein while a WM-to-AA substitution in the hexapeptide of Hoxa1 led to a drastic loss-of-function [Remacle et al., 2002]. Actually, this insertion disrupted the TFDWMK hexapeptide between the tryptophan and methionine residues, but the first inserted residue being a methionine, the critical WM dyad required for the interaction with Pbx is maintained. The TFDWMK sequence is thus splitted between the methionine and lysine residues, which according to structural data [Phelan et al., 1995; Passner et al., 1999; Piper et al., 1999] should not profoundly affect the Hox–Pbx interaction.

In conclusion, although Hoxa1 seems to be highly permissive to potentially destructuring mutations, pentapeptide insertion mutagenesis identified the homeodomain and both the N- and C-terminus of the protein as critical determinants for transcriptional activation (Fig. 8). The N-terminus and homeodomain of Hoxa1 are both necessary to stimulate transcription, whereas the serine–threonine–proline rich C-terminal motif exerts a modulatory influence on this

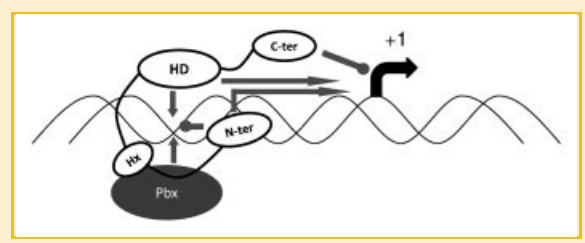


Fig. 8. Functional organization of Hoxa1. Hoxa1 interacts with Pbx proteins through its hexapeptide (Hx) and binds to DNA by its homeodomain (HD). The conserved N-terminal end of Hoxa1 (N-ter) exerts a regulatory influence on the cooperative DNA binding by Hoxa1 and Pbx. The homeodomain and N-terminus of the protein are together required to stimulate transcription initiation. The serine–threonine–proline rich C-terminus (C-ter) of the protein modulates its transcriptional activation potential.

activity. Finally, the N-terminus of Hoxa1 was also shown to modulate the synergistic interaction between Hoxa1 and Pbx1a during DNA–protein complex assembly.

ACKNOWLEDGMENTS

We are grateful to Michel Ghislain for his assistance in performing the yeast assays. We thank Bernard Peers for kindly providing the pCS2-Prep1 vector, Mark S. Featherstone for the *b1-ARE* reporter and Jin Chen and H. Earl Ruley for the *Epha2-r4-luciferase* reporter plasmid. This work was supported by the Belgian Fund for Scientific Research (FNRS, FRSM grant), the ''Direction Générale des Technologies, de la Recherche et de l'Energie'' of the Walloon Region (WALEOII grant), and the Fonds Spéciaux de Recherche (FSR) of the Université catholique de Louvain (UCL). B.L. held a FRIA fellowship from the FNRS and a FSR grant from UCL.

REFERENCES

Berthelsen J, Kilstrup-Nielsen C, Blasi F, Mavilio F, Zappavigna V. 1999. The subcellular localization of PBX1 and EXD proteins depends on nuclear import and export signals and is modulated by association with PREP1 and HTH. Genes Dev 13:946–953.

Cao Y, Hallet B, Sherratt DJ, Hayes F. 1997. Structure–function correlations in the XerD site-specific recombinase revealed by pentapeptide scanning mutagenesis. J Mol Biol 274:39–53.

Capecchi M. 1997. Hox genes and mammalian development. Cold Spring Harb Symp Quant Biol 62:273–281.

Chang CP, Brocchieri L, Shen WF, Largman C, Cleary ML. 1996. Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. Mol Cell Biol 16:1734–1745.

Chariot A, Castronovo V. 1996. Detection of HOXA1 expression in human breast cancer. Biochem Biophys Res Commun 222:292–297.

Chariot A, van Lint C, Chapelier M, Gielen J, Merville MP, Bours V. 1999a. CBP and histone deacetylase inhibition enhance the transactivation potential of the HOXB7 homeodomain-containing protein. Oncogene 18:4007–4014.

Chariot A, Princen F, Gielen J, Merville M-P, Franzoso G, Brown K, Siebenlist U, Bours V. 1999b. Ikappa B-alpha enhances transactivation by the HOXB7 homeodomain-containing protein. J Biol Chem 274:5318–5325.

Chen J, Ruley HE. 1998. An enhancer element in the EphA2 (Eck) gene sufficient for rhombomere-specific expression is activated by HOXA1 and HOXB1 homeobox proteins. J Biol Chem 273:24670–24675.

Chen DC, Yang BC, Kuo TT. 1992. One-step transformation of yeast in stationary phase. Curr Genet 21:83–84.

Chen Y, Knezevic V, Ervin V, Hutson R, Ward Y, Mackem S. 2004. Direct interaction with Hoxd proteins reverses Gli3-repressor function to promote digit formation downstream of Shh. Development 131:2339–2347.

Chisaka O, Musci TS, Capecchi MR. 1992. Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene Hox-1.6. Nature 355:516–520.

Cillo C, Cantile M, Faiella A, Boncinelli E. 2001. Homeobox genes in normal and malignant cells. J Cell Physiol 188:161–169.

Derossi D, Calvet S, Trembleau A, Brunissen A, Chassaing G, Prochiantz A. 1996. Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. J Biol Chem 271:18188–18193.

Di Rocco G, Mavilio F, Zappavigna V. 1997. Functional dissection of a transcriptionally active, target-specific Hox–Pbx complex. EMBO J 16:3644–3654.

Di Rocco G, Gavalas A, Popperl H, Krumlauf R, Mavilio F, Zappavigna V. 2001. The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function. J Biol Chem 276:20506–20515.

Dubnau J, Struhl G. 1996. RNA recognition and translational regulation by a homeodomain protein. Nature 379:694–699.

Ferretti E, Marshall H, Popperl H, Maconochie M, Krumlauf R, Blasi F. 2000. Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. Development 127:155–166.

Ferretti E, Cambronero F, Tumpel S, Longobardi E, Wiedemann LM, Blasi F, Krumlauf R. 2005. Hoxb1 enhancer and control of rhombomere 4 expression: Complex interplay between PREP1–PBX1–HOXB1 binding sites. Mol Cell Biol 25:8541–8552.

Gehring WJ, Qian YQ, Billeter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, Otting G, Wuthrich K. 1994. Homeodomain–DNA recognition. Cell 78:211–223.

Goudet G, Delhalle S, Biemar F, Martial JA, Peers B. 1999. Functional and cooperative interactions between the homeodomain PDX1, Pbx, and Prep1 factors on the somatostatin promoter. J Biol Chem 274:4067–4073.

Hallet B, Sherratt D, Hayes F. 1997. Pentapeptide scanning mutagenesis: Random insertion of a variable five amino acid cassette in a target protein. Nucleic Acids Res 25:1866–1867.

Hayes F, Hallet B. 2000. Pentapeptide scanning mutagenesis: Encouraging old proteins to execute unusual tricks. Trends Microbiol 8:571–577.

Jacobs Y, Schnabel CA, Cleary ML. 1999. Trimeric association of Hox and TALE homeodomain proteins mediates Hoxb2 hindbrain enhancer activity. Mol Cell Biol 19:5134–5142.

Kataoka K, Yoshitomo-Nakagawa K, Shioda S, Nishizawa M. 2001. A set of Hox proteins interact with the Maf oncoprotein to inhibit its DNA binding, transactivation, and transforming activities. J Biol Chem 276:819–826.

Kmita M, Duboule D. 2003. Organizing axes in time and space; 25 years of colinear tinkering. Science 301:331–333.

LaRonde-LeBlanc NA, Wolberger C. 2003. Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. Genes Dev 17:2060–2072.

Li X, Nie S, Chang C, Qiu T, Cao X. 2006. Smads oppose Hox transcriptional activities. Exp Cell Res 312:854–864.

Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. 2006. Intrinsic disorder in transcription factors. Biochemistry 45:6873–6888.

Liu Y, Matthews KS, Bondos SE. 2008. Multiple intrinsically disordered sequences alter DNA binding by the homeodomain of the *Drosophila* Hox protein Ultrabithorax. J Biol Chem 283:20874–20887.

Lufkin T, Dierich A, LeMeur M, Mark M, Chambon P. 1991. Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. Cell 66:1105–1119.

Lumsden A, Krumlauf R. 1996. Patterning the vertebrate neuraxis. Science 274:1109–1115.

Maizel A, Bensaude O, Prochiantz A, Joliot A. 1999. A short region of its homeodomain is necessary for engrailed nuclear export and secretion. Development 126:3183–3190.

Mann R, Affolter M. 1998. Hox proteins meet more partners. Curr Opin Genet Dev 8:423–429.

Matis C, Chomez P, Picard J, Rezsohazy R. 2001. Differential and opposed transcriptional effects of protein fusions containing the VP16 activation domain. FEBS Lett 499:92–96.

Merabet S, Kambris Z, Capovilla M, Bérenger H, Pradel J, Graba Y. 2003. The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. Dev Cell 4:761–768.

Moens CB, Selleri L. 2006. Hox cofactors in vertebrate development. Dev Biol 291:193–206.

Niessing D, Blanke S, Jäckle H. 2002. Bicoid associates with the 5'-cap-bound complex of caudal mRNA and represses translation. Genes Dev 16:2576–2582.

Passner JM, Ryoo HD, Shen L, Mann RS, Aggarwal AK. 1999. Structure of a DNA-bound Ultrabithorax–Extradenticle homeodomain complex. Nature 397:714–719.

Phelan ML, Featherstone MS. 1997. Distinct HOX N-terminal arm residues are responsible for specificity of DNA recognition by HOX monomers and HOX PBX heterodimers. J Biol Chem 272:8635–8643.

Phelan ML, Rambaldi I, Featherstone MS. 1995. Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. Mol Cell Biol 15:3989–3997.

Pinsonneault J, Florence B, Vaessin H, McGinnis W. 1997. A model for extradenticle function as a switch that changes HOX proteins from repressors to activators. EMBO J 16:2032–2042.

Piper DE, Batchelor AH, Chang C-P, Cleary ML, Wolberger C. 1999. Structure of a HoxB1-Pbx1 heterodimer bound to DNA: Role of the hexapeptide and a fourth homeodomain helix in complex formation. Cell 96:587–597.

Prince F, Katsuyama T, Oshima Y, Plaza S, Resendez-Perez D, Berry M, Kurata S, Gehring WJ. 2008. The YPWM motif links Antennapedia to the basal transcriptional machinery. Development 135:1669–1679.

Prochiantz A. 2000. Messenger proteins: Homeoproteins, TAT and others. Curr Opin Cell Biol 12:400–406.

Remacle S, Shaw-Jackson C, Matis C, Lampe X, Picard J, Rezsohazy R. 2002. Changing homeodomain residues 2 and 3 of Hoxa1 alters its activity in a cell-type and enhancer dependent manner. Nucleic Acids Res 30:2663–2668.

Remacle S, Abbas L, De Backer O, Pacico N, Gavalas A, Gofflot F, Picard JJ, Rezsohazy R. 2004. Loss of function but no gain of function caused by amino acid substitutions in the hexapeptide of Hoxa1 in vivo. Mol Cell Biol 24:8567–8575.

Rivera-Pomar R, Niessing D, Schmidt-Ott U, Gehring WJ, Jackle H. 1996. RNA binding and translational suppression by bicoid. Nature 379:746–749.

Ronshaugen M, McGinnis N, Mc Ginnis W. 2002. Hox protein mutation and macroevolution of the insect body plan. Nature 415:848–849.

Saleh M, Rambaldi I, Yang X-J, Featherstone MS. 2000. Cell signaling switches HOX–PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases. Mol Cell Biol 20:8623–8633.

Schild-Poulter C, Pope L, Giffin W, Kochan JC, Ngsee JK, Traykova-Andonova M, Hache RJ. 2001. The binding of Ku antigen to homeodomain proteins promotes their phosphorylation by DNA-dependent protein kinase. J Biol Chem 276:16848–16856.

Schnabel CA, Abate-Shen C. 1996. Repression by HoxA7 is mediated by the homeodomain and the modulatory action of its N-terminal-arm residues. Mol Cell Biol 16:2678–2688.

Shanmugam K, Green NC, Rambaldi I, Saragovi HU, Featherstone MS. 1999. PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. Mol Cell Biol 19:7577–7588.

Shen W-F, Krishnan K, Lawrence HJ, Largman C. 2001. The HOX homeodomain proteins block CBP histone acetyltransferase activity. Mol Cell Biol 21:7509–7522.

Shen W, Chrobak D, Krishnan K, Lawrence HJ, Largman C. 2004. HOXB6 protein is bound to CREB-binding protein and represses globin expression in a DNA binding-dependent, PBX interaction-independent process. J Biol Chem 279:39895–39904.

Studer M, Gavalas A, Marshall H, Ariza-McNaughton L, Rijli FM, Chambon P, Krumlauf R. 1998. Genetic interactions between Hoxa1 and Hoxb1 reveal new roles in regulation of early hindbrain patterning. Development 125:1025–1036.

Tour E, Hittinger CT, McGinnis W. 2005. Evolutionarily conserved domains required for activation and repression functions of the *Drosophila* Hox protein Ultrabithorax. Development 132:5271–5281.

Vigano MA, Di Rocco G, Zappavigna V, Mavilio F. 1998. Definition of the transcriptional activation domains of three human HOX proteins depends on the DNA-binding context. Mol Cell Biol 18:6201–6212.

Williams TM, Williams ME, Heaton JH, Gelehrter TD, Innis JW. 2005. Group 13 HOX proteins interact with the MH2 domain of R-Smads and modulate Smad transcriptional activation functions independent of HOX DNA-binding capability. Nucleic Acids Res 33:4475–4484.

Wolberger C. 1996. Homeodomain interactions. Curr Opin Struct Biol 6:62–68.

Zappavigna V, Sartori D, 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, Bianchi ME. 1996. HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. EMBO J 15:4981–4991.

Zhang X, Zhu T, Chen Y, Mertani HC, Lee KO, Lobie PE. 2003. Human growth hormone-regulated HOXA1 is a human mammary epithelial oncogene. J Biol Chem 278:7580–7590.

Zhao JJ, Lazzarini RA, Pick L. 1996. Functional dissection of the mouse Hox-a5 gene. EMBO J 15:1313–1322.