INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Université d'Ottawa · University of Ottawa

FUNCTIONAL ANALYSIS OF DLX HOMEODOMAIN PROTEINS FROM THE ZEBRAFISH

By

Nathalie Chartrand

A thesis submitted to the School of Graduate Studies and Research in partial fulfillment for the degree

of

Master of Science

Department of Anatomy and Neurobiology

University of Ottawa

Ottawa, Ontario

August 17, 1999

Acknowledgments/ Remerciements

Je voudrais tout d'abord remercier mon superviseur Marc Ekker de m'avoir accueillie dans son laboratoire. Je voudrais le remercier de m'avoir bien guidée dans mon projet. Je veux le remercier de sa patience et de ses encouragements. Je voudrais aussi le remercier de m'avoir enseigné comment penser et aussi de m'avoir permis d'apprendre à travailler de façon indépendante.

I would like also to thank the members of my committee, Marie-Andrée Akimenko, William Staynes and Johne Liu for their good advices.

I would like also to thank the members of the laboratory, Liz Quint who helped me with the confocal microscopy and her encouragements, Gary Hatch for his support, Quioming Long, Keon Park, Jennifer Miles for her moral support and her comprehension, Lynda Laforest pour son aide technique et sa grande écoute. I would like to thank Genny Giroux who first thought me the basic techniques in molecular biology. I would like also to thank Cricho Brown and German Poleo for the nice discussions we had and Jun Ming.

A big thank you to Gratien Préfontaine who helped me with the western analysis and also with the computers.

Un merci spécial à Lucille Joly, mon amie, qui m'a écoutée quand j'en avais le plus besoin. Je te remercie pour ta gentillesse, ta compréhension et tes encouragements qui m'on aidée à continuer. J'apprécie énormément ton amitié.

A very special thank to my dear friend Ted Zerucha who helped me a lot with my project and support me. I thank you for your comprehension when I needed the most. Thank you for your encouragements and comfort me when times were difficult. Thank you also for helping me getting more confidence in myself.

Finallement un énorme merci à ma famille, tout spécialement ma mère à qui je dois beaucoup. Je te remercie de m'avoir soutenue et encouragée quand j'en avais besoin. Merci de ta patience envers moi. Tu m'as toujours écoutée et réconfortée quand mon moral était bas. Merci pour ton amour qui m'est si cher. Je te remercie d'avoir été là avec moi.

Abstract:

The *dlx* homeobox genes of vertebrates are transcriptional regulators involved in the development of the ventral forebrain, otic vesicle and inner ear, visceral arches, olfactory placodes and fins/limbs. The zebrafish genome contains eight *dlx* genes whose expression patterns during development are partially overlapping. Experiments in cultured cells and transgenic zebrafish embryos were used to understand the functional specificity and the biochemical properties of Dlx proteins. Previous work by us and others showed that Dlx proteins can act as transcriptional activators. In order to determine how the various Dlx proteins differ from each other, transfection experiments in cultured cells were done to identify the Dlx transcriptional activation domains. I found that an activation domain was located in the N-terminal region of Dlx1, Dlx2 and Dlx4. In addition, to test the hypothesis that the partially overlapping expression of *dlx* genes during development is the result, in part, of specific cross-regulatory interactions between the genes, overexpression of mutant versions of the *dlx* genes into transgenic zebrafish embryos were carried out. Injection into one cell embryo of synthetic mRNA coding for a chimeric protein including the amino-terminal region of Dlx3 and the carboxy-terminal half of Dlx2, resulted in a loss or decrease of endogenous *dlx4* expression in the visceral arches, the otic vesicles and the forebrain of 27 hours embryos. Furthermore, I demonstrated that overexpression of that chimera, which disturbed the endogenous *dlx4* expression, results in malformations of the craniofacial cartilages. It suggests that Dlx3 is normally a positive regulator of *dlx4* expression and that the chimera Dlx3-Dlx2 protein interferes with the function of the endogenous Dlx3. These results suggest that the

combinatorial patterns of *dlx* expression are due at least in part to cross-regulatory interactions between the different *dlx* genes.

Résumé:

Les gènes à boîte homéo des vertébrés sont des régulateurs de la transcription impliqués dans le développement du cerveau ventral, des vésicules otiques et de l'oreille interne, des arcs viscéraux, des placodes olfactives et des nagoires/membres. Le génome du poisson zèbre contient huit gènes *dlx* dont le pattern d'expression est combinatoire. Des expériences dans des cellules en culture ainsi que la production d'embryons transgéniques de poisson zèbres ont été effectuées dans le but de mieux comprendre le rôle spécifique et les propriétés biochimiques des protéines Dlx. Des expériences menées par nous et par d'autres laboratoires démontrent que les protéines Dlx sont des activateurs de la transcription. Dans le but de déterminer comment les Dlx diffèrent les uns des autres, d'autres expériences de transfections dans des cellules en culture ont été effectuées pour identifier les domaines d'activation des Dlx. J'ai trouvé un domaine d'activation localisé dans la partie N-terminale de Dlx1, Dlx2 et Dlx4. De plus, pour analyser l'hypothèse voulant que l'expression combinatoire des *dlx* soit le résultat, en partie, de la régulation croisée entre les Dlx, des versions mutantes de *dlx3* ont été injecté dans des embryons au stade une cellule. L'injection d'ARNm synthétiques codant pour une protéines composée de la région N-terminale de Dlx3 ainsi que de la région C-terminale de Dlx2 a engendré une perte ou une diminution de l'expression endogène de

dlx4, dans les arcs viscéraux, les vésicules otiques et le cerveau ventral. De plus je démontre que la surexpression de la chimère, qui influence l'expression de *dlx4*, produit des malformations des cartilages de la stucture craniofaciale des poissons. Les résultats obtenus suggèrent que Dlx3 est un régulateur de la transcription de *dlx4* et que la chimère Dlx3-Dlx2 interfère avec la fonction de Dlx3 endogène. Ces résultats suggèrent que l'expression combinatoire des *dlx* est causée, en partie, par de la régulation croisée entre les différents gènes *dlx*.

Table of contents

1- Introduction:

2. Material and Methods

3. Results

4.Discussion

List of Figures

List of tables

List of abreviations

aa	amino acid
ala	alanine
Antp	*antennapedia* gene
asp	aspartic acid
cAMP	cyclic adenosine monophosphate
CAT	chloramphenicol acetyl-transferase
CDNA	complementary DNA
CHO	chinese hamster ovary cell
C-terminal	carboxy-terminal
Dll	Distal-less gene (*Drosophila*)
Dlx	Distal-less gene (Vertebrates)
DNA	Deoxyribonucleic acid
dpf	days post-fertilization
en	engrailed gene
Ftz	fushi tarazu gene
Gal-4 DBD	Gal-4 DNA Binding Domain
HD	homeodomain
Hox	clustered homeobox genes

hpf	hours post-fertilization
M	molarity
mRNA	messenger ribonucleic acid
msx	muscle segmented homeobox gene (vertebrate)
M.W.	Molecular Weight
NMR	nuclear magnetic resonance
N-terminal	amino-terminal
PCR	polymerase chain reaction
PFA	paraformaldehyde
PKA	protein kinase A
R.F.	RNase free
T	threonine

1. Introduction

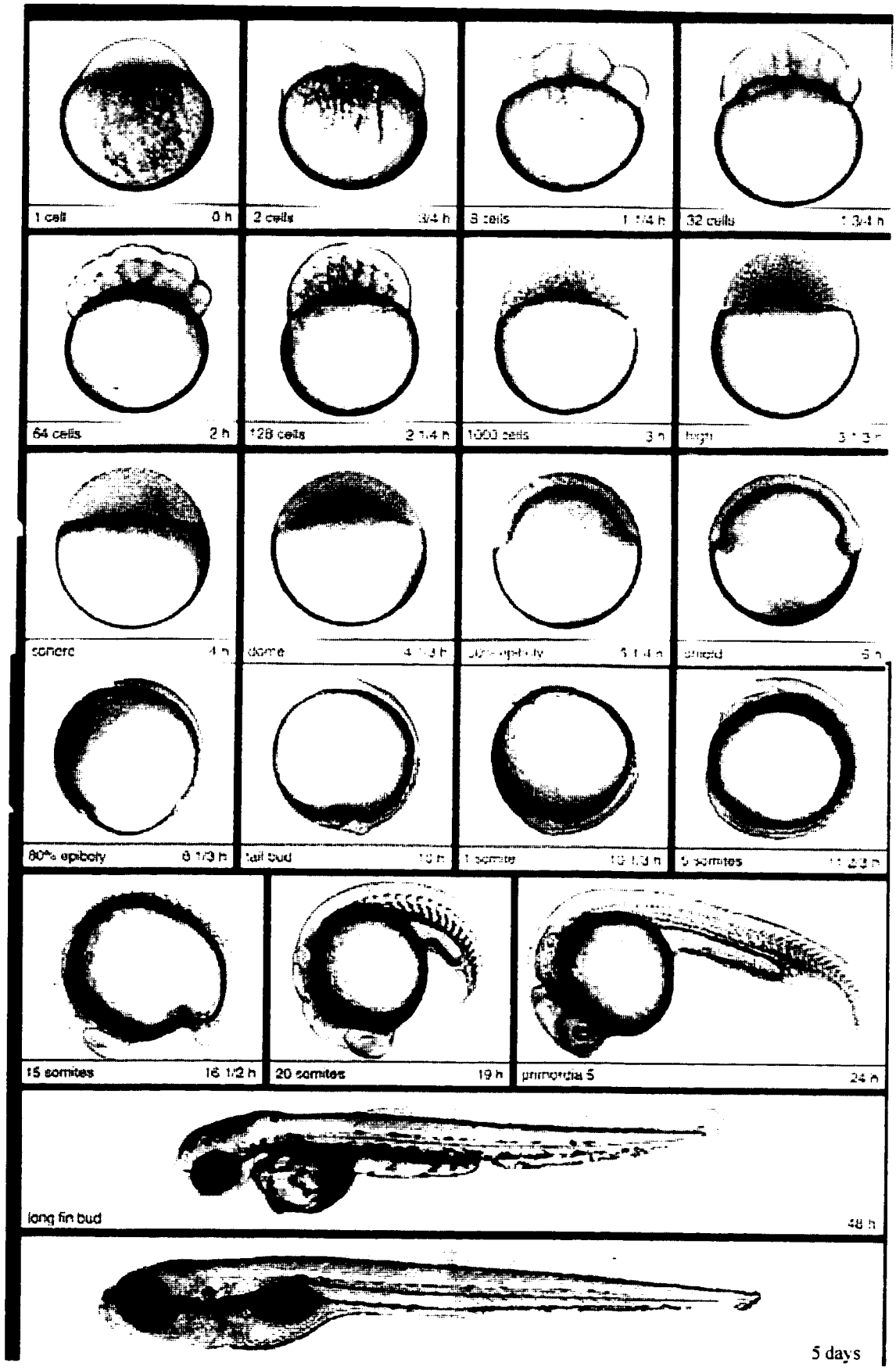
1.1. Zebrafish as a developmental model.

The genetics of invertebrate embryonic pattern formation is traditionally studied using the *Drosophila* developmental model, whereas for vertebrates the mouse is the most extensively studied model. However it is not easy to work with the mouse, because the embryos develop inside the female, only 10 to 14 embryos are produced per mother and the gestation time is long; 3 weeks. It was of great importance to find another model for vertebrate development that would be easier to work with. The zebrafish is a very good model since its development enables the use of both experimental embryology and molecular genetic methodologies.

More specifically zebrafish have distinct features, which make it a preferred vertebrate developmental model. They are small (3-4cm long) tropical fresh water fish, easy to maintain in large numbers. Embryonic development is very rapid; after only 12 hours from the time of fertilization, one can visualize the establishment of a body plan that is typically vertebrate (Kimmel., 1989) (see figure 1-1 for developmental stages). They reach sexual maturity within 3 months. Females lay anywhere from 300 to 500 eggs per spawning therefore facilitating genetic analyses. Eggs are fertilized externally and the embryos are lucent, facilitating the observation of individual cells during development. The translucent embryo allows one to view the cell movements during gastrulation, the formation of the brain, and the beating heart. The zebrafish is a good model to use for genetic studies, because of its nice features. The embryos are available at any stage for

Figure 1-1: Zebrafish development

Developmental stages of a zebrafish embryo from 1 cell stage (0hpf) to a swimming larva (5dpf). Dorsal is to the top, anterior to the left.


1 cell 0 h
2 cells
8 cells
32 cells
64 cells 2 h
128 cells 2 1/4 h
1000 cells 3 h
high
sphere 4 h
dome
shield
80% epiboly 8 1/3 h
tail bud
1 somite
5 somites
15 somites 16 1/2 h
20 somites 19 h
primordia 5 24 h
long fin bud 48 h
5 days

screening and analysis. Developmental changes due to mutations, even those that lead to death of the embryos, can be studied with relative ease.

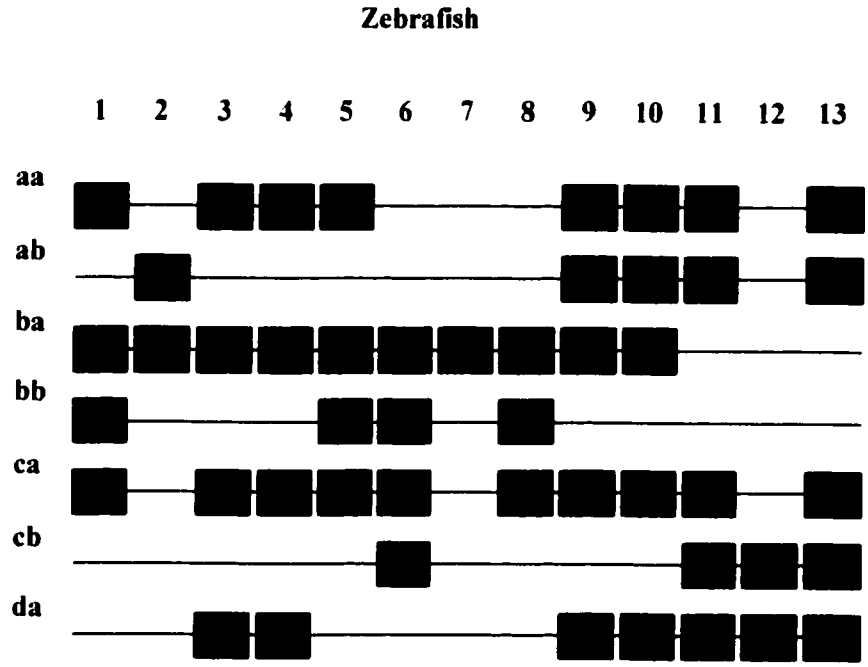
1.2. Homeodomain proteins during development.

Homeodomain proteins comprise a large class of transcription factors that have been shown to be essential regulators of many developmental processes, ranging from organization of the basic body plan to terminal differentiation of individual tissues by coordinating the expression of specific sets of target genes (McGinnis and Krumlauf., 1992). The homedomain is a highly conserved 60 amino acid element that is responsible for sequence-specific interaction with DNA.

The most famous family of homeodomain proteins are the Hox proteins. In vertebrates, such as the mouse, the *Hox* genes are clustered in four genomic loci containing 9-11 genes (McGinnis and Krumlauf. 1992), whereas, in zebrafish seven clusters have been identified and each one of them was found to contain 4 to 10 genes (Amores et al. 1998) (Figure 1-2). These genes encode transcription factors which control cell fate and developmental patterns in all metazoans, leading to the generation of morphological differences along the body axes. This class of protein was first discovered by studying spontaneous mutations found in *Drosophila*, termed homeotic phenotypes. Homeosis refers to the substitution of one body part for another. The first homeotic mutation, discovered by Bridges in 1915, resulted in a partial or complete segmental transformation. (Gehring., 1994). Although homeobox genes have been often equated with *Hox* genes, not all homeobox genes are *Hox* genes. Indeed, each species may have up to several hundred homeobox genes as part of their genome. Homeobox genes have been classified into families on the basis of the amino acid sequence of their homeodomain.

Figure 1-2: Zebrafish Hox clusters

Schematic representation of the zebrafish *Hox* clusters. Each cluster is located on a different chromosome and contains 4 to 10 genes which correspond to the black boxes


Zebrafish
1
2
3
4
5
6
7
8
9
10
11
12
13
aa
ab
ba
bb
ca
cb
da

The *distal- less / dlx* genes, the object of this work, constitute one such family The characteristics of homeodomain proteins as transcription factors such as the structure of their homeodomain, the DNA binding specificity, and their activation domains are going to be discussed in this introduction. We will then present the background that pertains to the distal-less family of homeobox genes.

1.3. Homeobox genes as transcription factors.

1.3.1. Different types of DNA binding domains.

The homeobox was first identified in two homeotic genes of *Drosophila*, *Antennapedia* and *Ultrabithorax* (Gehring, 1987; McGinnis et al., 1984; Scott et al., 1989; Scott and Weiner, 1984). This motif has been well conserved through evolution and is found in a wide variety of eukaryotic genes, including invertebrate, vertebrate and yeast genes that encode transcriptional regulators (Laughon and Scott, 1984; Pabo and Sauer, 1984; Scott et al., 1989).

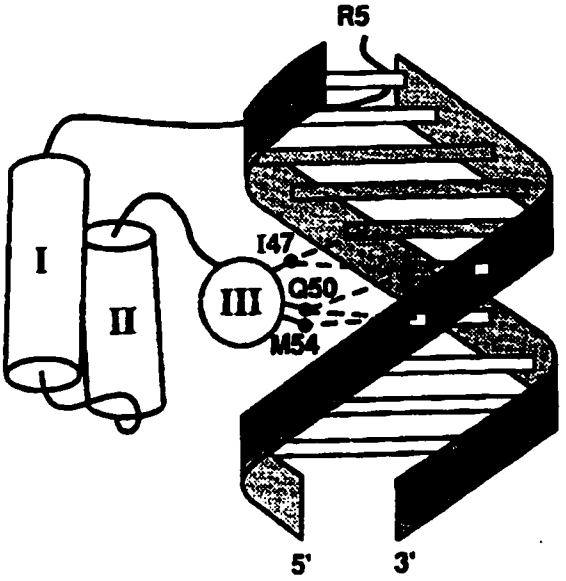
The homeobox genes are transcription factors for which very few of the downstream targets during development have been identified. Only phenotypic studies of mutated animals gave insights of what is the possible role of the *Hox* genes. It has been also possible to find that the *Hox* genes are able to regulate themselves as well as other *Hox* genes. However the molecular mechanisms by which they exert their effect are not known. Two groups hypothesised that the homeodomain was the DNA-binding domain. The rationale behind this is the observation of sequence similarity of the homeodomain to the helix-turn-helix motif of prokaryotic gene regulatory proteins (Laughon and Scott, 1984; Shepherd et al., 1984) (Figure 1-3 A), as well as the similarity to the mating type proteins MATa1 and α2 in yeast (Shepherd et al., 1984) which are also regulators of

Figure 1-3: Schematic showing the homeodomain-DNA complexes of Antp and the structure of the POU domain of Oct-1 bound to DNA.

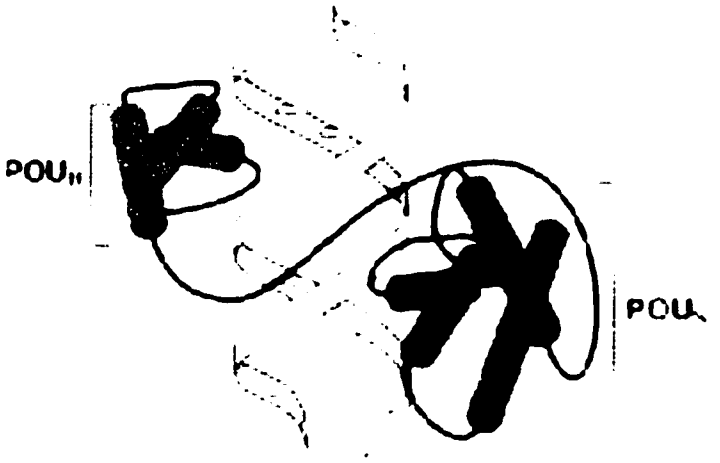
(A) View along the axis of the recognition helix (III) of the Antp homeodomain. Amino acid residues that establish contacts to specific bases are indicated. For residues in the recognition helix, these contacts, which are indicated by dotted lines, are in the major groove; for those in the N-terminal arm, they are in the minor groove of the DNA. The TAAT core motif is shaded (Gehring et al. 1994).

(B) Diagrammatic representation of the crystal structure of the Oct-1 POU domain monomer bound to an octamer element (ATGCAAAT on the lower strand). The Oct-1 POU-specific and POU homeodomain bind to opposite faces of the DNA and have a tail-to-tail orientation. The positions of the Oct-1 POU-specific domain (POU_S) and POU homeodomain (POU_H) on the ATGC and AAAT half-sites, respectively, are diagramed on the sequence on the next page (Ryan et al. 1997).

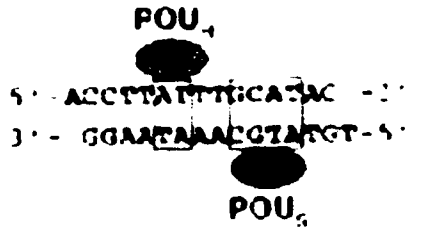
A



B



OCT-1



transcription. This hypothesis was verified as homeodomain proteins were demonstrated to bind DNA (Beachy et al., 1988: Desplan et al., 1985; Desplan et al., 1988; Fainsod et al., 1986; Hoey and Levine, 1988; Laughon et al., 1988; Muller et al., 1988). Later, the homeodomain was shown to be sufficient for DNA-binding, the structure of the homeodomain was determined and the ability of the homeodomain proteins to regulate transcription was demonstrated (Jaynes and O'Farrel., 1988; Han et al.., 1989; Krasnow et al., 1989; Gehring et al., 1994).

The ability of homeodomain proteins to bind DNA was demonstrated by a number of groups in the 3-5 years following the initial reports describing the discovery of the homeobox (Beachy et al., 1988; Desplan et al., 1985; Desplan et al., 1988; Fainsod et al., 1986; Hoey and Levine, 1988; Laughon et al., 1988; Muller et al., 1988). One of the earlier studies identified a nucleotide sequence fragment from upstream the murine Hoxa3 gene that was bound by nuclear extracts from mouse embryonic cells (Fainsod et al., 1986). To determine if one of the proteins, Hoxa3 which was likely present in those extracts based on its embryonic expression pattern, could bind this sequence, Hoxa3 was expressed in bacteria and subsequently shown to bind that fragment. A similar study assayed the ability of recombinant expressed engrailed (*en*), a homeodomain-containing protein from *Drosophila* but not a member of the Hox proteins, to bind a short repeated nucleotide sequence motif, TCAATTAAAT, found in a known element responsible for regulating *en* (Desplan et al., 1988). Full length recombinant *en* was demonstrated to bind this sequence. In addition, by making use of deletions of the *en* protein, the authors also demonstrated that the homeodomain alone was sufficient for binding to this sequence. Another distantly related homeodomain transcription factor, Fushi tarazu (Ftz) was found to bind also to this core sequence.

Another family of transcription factors involved in regulating key developmental processes during development, is characterized by the POU DNA-binding motif. The POU domain family of transcription factors was defined following the observation that the products of three mammalian genes, *Pit-1, Oct-1 and Oct-2* and the protein encoded by the *Caenorhabditis elegans* gene *unc-86* shared a region of homology, known as the POU domain (Bodner et al. 1988; Clerc et al. 1988; Finney et al. 1988; Herr et al. 1988). The POU domain is a bipartite DNA-binding domain (Sturm and Herr 1988; Ingraham et al. 1990; Botfield et al. 1992; Verrijzer et al. 1992), consisting of two highly conserved regions tethered by a variable linker. The ~75-amino acid amino-terminal region was called the POU specific domain and the carboxy-terminal 60 amino acid region, the POU homeodomain.

High-affinity site-specific DNA binding by POU domain transcription factors requires both the POU-specific domain and the POU homeodomain (Sturm and Herr, 1988; Ingraham et al. 1990; Kristie and Sharp 1990). The two subdomains can cooperatively bind DNA even when they are not joined by a linker (Klemm and Pabo 1996). Resolution of the crystal structures of Oct-1 and Pit-1 POU domains and nuclear magnetic resonance (NMR) studies showed that the POU-specific domain consisted of 4 α helices, with the second and third helices forming a structure similar to the helix-turn-helix motif of the λ and 434 repressors; several of the DNA base contacts are also conserved (Klemm et al. 1994). The POU homeodomain structure is similar to the engrailed, Matα2 and Antennapedia homeodomains (Quian et al., 1989; Kissinger et al., 1990; Wolberger et al., 1991) The POU-specific domain binds the 5' ATGC portion of an octamer element and the POU homeodomain binds the 3' A/T-rich portion of the site (Klemm and Pabo 1996) (Figure 1-3 B).

1.3.2. Specificity of homeodomain-DNA interaction

The specificity of action that each homeodomain protein exhibits *in vivo*, however, has not yet been duplicated *in vitro*. In fact, homeodomain proteins as mentioned above tend to bind short A/T-rich consensus sites that are often indistinguishable from one another (Hayashi and Scott, 1990; Kalionis and O'Farrell, 1993; Gehring et al., 1994).

DNA binding to specific recognition sequences in the promoters of target genes mediates many activities of homeodomain-containing proteins (Hanes et al., 1989; Schier et al., 1993). X-ray crystallographic analysis of protein-DNA complexes involving the homeodomains of *Drosophila* engrailed (Kissinger et al., 1990) and yeast Matα2 (Wolberger et al., 1991), and nuclear magnetic resonance studies of the Antennapedia (Antp) homeodomain-DNA complex (Otting et al., 1990), have revealed that two regions of the homeodomain make base-specific contacts with DNA. The third α helix lies along the major groove while the N-terminal arm (a small sequence located at the N-terminal end of the homeodomain), reaches behind the phosphate backbone to contact bases via the minor groove (Figure 1-3 A). Although engrailed and Matα2 homeodomains exhibit only 27% sequence identity, their DNA binding conformations are superimposable (Wolberger et al., 1991). A number of phosphate contacts made by conserved residues and a base-specific contact made by the invariant asparagine at homeodomain position 51 (Asn-51) indicate that even diverged homeodomains bind DNA in similar fashions.

The importance of the amino-terminal arm in conferring binding site specificity outside the ATTA/TAAT core element has been reported by a number of groups. One of the first studies to examine this elegantly compared the amino-terminal arms of the murine *Hoxd4* and *Hoxa1* genes (Phelan et al., 1994). The *Hoxd4* gene has an

autoregulatory function mediated through its binding a relatively well characterized element in its promoter. Hoxa1, though, is unable to bind and activate transcription through this same region. However, when the second and third amino acids of the Hoxa1 amino-terminal arm were converted to those corresponding positions of Hoxd4, the Hoxa1 mutant bound and activated transcription through the *Hoxd4* autoregulatory motif. Based on what is known structurally of the homeodomain-DNA interaction, the amino acids at positions 2 and 3 of the amino-terminal arm contact the minor groove of the DNA double helix outside the ATTA/TAAT core sequence. Amino acids located further towards the carboxy-terminus such as at position 5 seem to impart some specificity within the core sequence as variations in that amino acid result in altered preferences for it.

Another potential source of specificity is through interactions with other DNA-binding transcription factors. Indeed a number of DNA-binding partners and cofactors have been identified in the past few years (Mann and Chan, 1996). Some of these partners change DNA-binding specificity whereas others alter the ability to activate or repress adjoining promoters. PBX1 is such an example. This gene participates in regulating normal differentiation by functioning in concert with the Hox proteins to activate or repress target gene expression (Rauskolb et al., 1993). Lu et al demonstrated that Hox proteins exhibit cooperative binding to ATCAATCAA with PBX1a and PBX1b as well as their oncogenic derivatives E2A-PBX1a and E2A-PBX1b (Lu et al., 1995). They showed that Hoxa-5, Hoxb-7, Hoxb-8, Hoxc-8 and Hoxd-4 exhibited cooperative binding to the PBX PRS (Pbx1-responsive sequence) with PBX1 proteins. This suggests that each protein binds a DNA sequence within the PRS in a specific manner. To determine if this cooperative binding between PBX1 and Hox proteins influence transcription, they performed cotransfection assays with a cytomegalovirus expression constructs encoding PBX1 proteins and Hoxb-8 or Hoxc-8. They observed that Hoxb-8 was able to suppress transactivation of E2A-PBX1 but did not inhibit general transcription, because the level

of E2A-PBX1 or PBX1 expressed from cotransfected vectors remained strong. These results provided the first evidence that expression of Hox proteins can have an impact on transcriptional regulation by PBX1 proteins.

Furthermore, Pöpperl et al demonstrated that rhombomere 4 expression is mediated by a positive and direct autoregulatory feedback mechanism that is cofactor dependent, and they provide *in vivo* and *in vitro* evidence that pbx proteins can act as the cofactors and cooperate with Hoxb-1 in autoregulation (Pöpperl et al. 1995).

Another potential source of specificity is the covalent addition of phosphate groups. Phosphorylation has been shown to affect a variety of transcription factor properties including structure, sub-cellular localization, DNA-binding affinity and specificity and the ability to activate transcription (Hunter and Karin, 1992; Karin, 1994; Hill and Treisman, 1995). All homeodomain proteins examined thus far are phosphorylated (Gay et al., 1988a; Krause et al., 1988; Krause and Gehring, 1989; Ronchi et al., 1993; Bourbon et al., 1995; Jaffe et al., 1997; Zwilling et al., 1997), indicating a likelihood for similar types of phosphorylation-induced changes in activity.

Experiments were done to examine the importance of phosphorylation on homeodomain protein function and specificity. Dong et al used *Fushi tarazu* expressed in *Drosophila* as a model to study phosphorylation (Dong et al., 1998). *Ftz* is expressed during 3 stages of embryogenesis: first in alternating segmental primordia, then in differentiating neurons in the central nervous system (CNS), and finally in portions of the gut and posterior epidermis (Hafen et al., 1984; Carroll and Scott, 1985; Krause et al., 1988). In this study they report that Ftz residue threonine 263 (T263), which is in the N-terminal arm of the homeodomain, is phosphorylated *in vitro* in a cAMP-dependent fashion by *Drosophila* embryo extracts and by purified cAMP-dependent protein kinase

(PKA). To test whether this phosphorylation event is required for Ftz activity, they mutated T263 to Ala and Asp to mimic the unphosphorylated and constitutively phosphorylated states, and subcloned the mutant DNAs into a 10 Kb fragment of Ftz genomic DNA, capable of rescuing Ftz$^-$ embryos to adulthood (Hiromi et al., 1985). The reconstituted genes were transformed into flies by P-element-mediated germline transformation and tested for their ability to rescue *Ftz* mutant embryos. They observed that the Ftz T263 mutated to an Ala failed to rescue *Ftz*$^-$ flies to adulthood whereas the T263 mutated for an Asp rescued the mutated flies. They showed with these results that T263 is phosphorylated when Ftz is in its active form, and that phosphorylation of T263 probably affects a protein-protein interaction.

Phosphorylation outside the homeodomain seems to be another mechanism that would affect the functions of homeodomain transcription factors such as subcellular localization, DNA binding and transcriptional activation. Bourbon et al showed that engrailed protein was able to be phosphorylated by a kinase named casein kinase II. They identified a region that allowed phosphorylation, which is about 50 residues to the N-terminal side of the homeodomain. They showed by *in vitro* studies, that phosphorylation of a truncated *en* enhances DNA binding. They suggested that the conformation of this protein may change after phosphorylation favouring DNA binding. (Bourbon et al., 1995)

Another protein has been shown to be phosphorylated *in vitro*. The protein in question here is Ultrabithorax which is involved in specifying unique features that distinguish the individual segments in *Drosophila*. This protein produces isoforms which are also phosphorylated. Gavis and Hogness reported 2 regions in the N-terminal that contains serine and threonine residues that can be phosphorylated. They had not elucidated the functional significance of this event. (Gavis and Hogness, 1991).

1.3.3. Activation domains of homeodomain proteins.

An activation domain is a sequence in a protein that allows protein-protein interactions necessary for gene transcription initiation. Its location in the homeodomain proteins has not been studied extensively. Among the few reports on that subject, one is about the Hox activation domains. Viganò et al cotransfected full length and deleted Hoxd-9, Hoxb-1 and Hoxb-3 with a reporter luciferase constuct. They studied the activation domains in different contexts which are: homologous context i.e. within a Hox protein binding either as a monomer to an ATTA-containing sequence or as a Hox-PBX dimer on a TGAT(T/G)NAT-containing sequence, or in a heterologous context after translocation to the yeast Gal-4 DNA binding domain (Gal-4 DBD) (Viganò et al., 1998) (see also figure 2-5). Activation domains have been identified in different regions of the Hox proteins depending on the context they have been studied. An activation domain was found to spread a large region of the Hoxd-9 from amino acid 75-222 in a homologous context. In contrast, in the heterologous context a Hoxd-9 activation domain was found to be located in the first 75 amino acids. Similarly, an Hoxb-1 activation domain was found between amino acids 38 and 90 when studied in a dimer context with PBX1, while in a heterologous context, (fused to Gal-4 DBD) an activation domain was found to spread to the entire N-terminal region. Finally, Hoxb3 activity was studied in a heterologous context, where transcriptional activation was promoted by the N-terminus and the C-terminus. They studied the activity of those regions in a heterologous context as well. They observed transcriptional activation promoted by the C-terminus only. Their data indicate that identification of transcriptionaly active regions is highly dependent on the context.

Vertebrate genes with homeoboxes similar in sequence to that of the *Drosophila distal-less* gene (Cohen et al. 1989 and Vachon et al 1992) constitute the *dlx* family. Vertebrate *dlx* genes include those of the mouse (Porteus et al., 1991; Price et al., 1991;

Robinson et al 1991), human (Simeone et al., 1994; Scherer et al 1995), rat (Zhao et al 1994), chick (Ferrari et al 1995), newt (Beauchemin and Savard, 1992), *Xenopus* (Asano et al 1992; Papalopulu and Kinter, 1993), axolotl (Mullen et al 1996) and zebrafish (Ekker et al., 1992; Akimenko et al., 1994; Stock et al., 1996). Most of our understanding of vertebrate *dlx* gene function during development comes from studies describing their embryonic expression patterns which are going to be discussed below.

1.4. The distal-less family of homeobox genes

1.4.1. Genomic organization

The vertebrate *Dlx* family consists of 6 or 8 genes in zebrafish. They are organized into 6 paralogous groups based on sequence similarities of the homeodomain which they encode (table 1-1). The *Dlx* genes can also be grouped into one of two clades based on homeodomain sequence as well as sequence similarities in the carboxy terminus (Stock et al., 1996). One clade is made up of the *Dlx1*, *Dlx6* and *Dlx7* paralogous groups and the other, the *Dlx2*, *Dlx3* and *Dlx5* paralogous groups. The proteins encoded by members of the clade of *Dlx2*, *Dlx3* and *Dlx5* (the last is called *dlx4* in zebrafish) also contain a short highly conserved amino acid sequence in the amino terminus (Akimenko et al. 1994; Stock et al. 1996). In addition, at least 6 of the known *Dlx* genes are organized as inverted convergently transcribed gene pairs,with each gene pair containing one gene from each clade (figure 1-4 right panel) . Thus *Dlx1* and *Dlx2* are linked pair as are *Dlx5* (*dlx4* in zebrafish) and *Dlx6* as well as *Dlx3* and *Dlx7* (Ozcelik et al., 1992; Simeone et al., 1994; McGuinness et al., 1996; Nakamura et al., 1996; Ellies et al., 1997a). Thus *Dlx1* and *Dlx2* are linked pair as are *Dlx5* (*dlx4* in zebrafish) and *Dlx6* as well as *Dlx3* and *Dlx7* (Ozcelik et al., 1992; Simeone et al., 1994; McGuinness et al., 1996; Nakamura et al., 1996; Ellies

et al., 1997a). As mentioned above, zebrafish has two additional *dlx* genes, *dlx5* and *dlx8*, which are not linked and that do not appear to exist in other vertebrates (Stock et al., 1996). The genomic organization of the Distal-less related genes is very well conserved from lower vertebrates to higher vertebrates. Each gene contains 3 exons, and the homeodomain spans exons 2 and 3.

1.4.2. *Distal-less evolution*

The vertebrate *Hox* genes are thought to arise from an ancestral *Hox* cluster that went through a series of gene duplication events (Stock et al 1996). A number of other gene families have members located on multiple *Hox*-containing chromosomes in vertebrates, suggesting that the *Hox* cluster duplications may have involved entire chromosomes or genomes (Ruddle et al., 1994a: Ruddle et al., 1994b). The *Distal-less* group is one of the family where the genes are linked to *Hox* clusters. The genomic organization and linkage to the *Hox* genes of the *Dll-related* gene family allows the development of a scenario by which this family likely evolved. Like the *Hox* genes the ancestral *distal-less* gene would have undergone an inversion/duplication event, thus creating two primordial vertebrate *dlx* genes A and B (Figure 1-4). This vertebrate *dlx* gene pair would have then undergone duplication events, thus creating the six inverted *Dlx* genes which are genomically linked (figure 1-4).

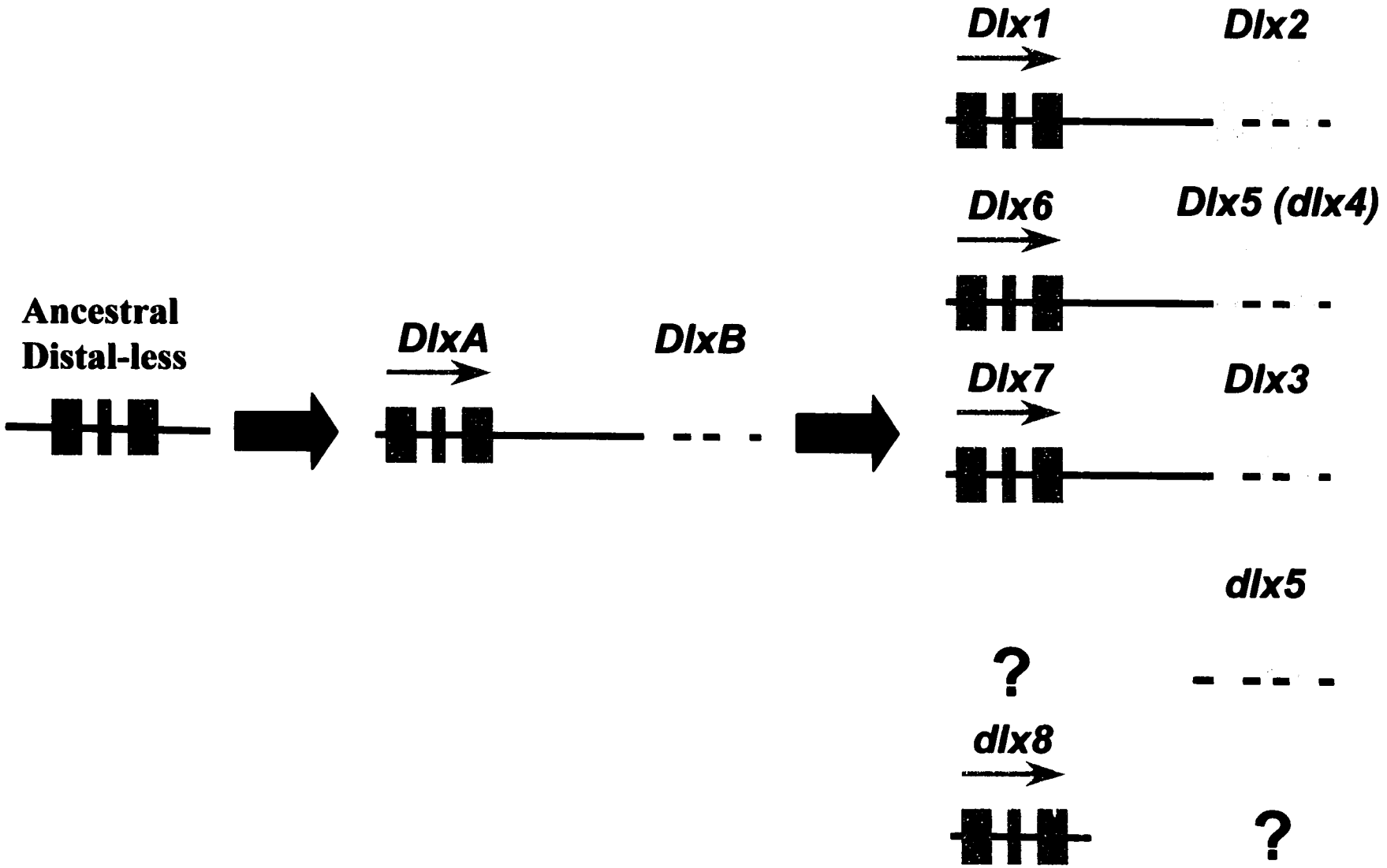
Table: 1- 1: *Dlx* orthologous groups

In vertebrate like the mouse, 6 *Dlx* genes have been found whereas in zebrafish 8 have been identified. Note that the two supplementary *dlx* genes in zebrafish which are *dlx5* and *dlx8* are members of the *Dlx2* and *Dlx7* groups respectively.

Dlx1	*Dlx6*	*Dlx7*
Human *DLX1*	Human *DLX6*	Human *DLX7*
Mouse *Dlx1*	Mouse *Dlx6*	Mouse *Dlx7*
Zebrafish *dlx1*	Zebrafish *dlx6*	Zebrafish *dlx7*
E. coqui EcDlx1	*Xenopus Xdll*	Zebrafish *dlx8*
		Newt *NvHBox5*
Dlx2	*Dlx5*	*Dlx3*
Human *DLX2*	Human *DLX5*	Human *DLX3*
Mouse *Dlx2*	Mouse *Dlx5*	Mouse *Dlx3*
Zebrafish *dlx2*	Zebrafish *dlx4*	Zebrafish *dlx3*
Zebrafish *dlx5*	*Xenopus X-dll3*	*Xenopus X-dll2*
Xenopus X-DLL1	Chicken *Dlx5*	*Xenopus Xdll-2*
Xenopus X-dll4	Rat *rDlx*	Newt *NvHbox4*
E. coqui EcDlx2	*E. coqui EcDlx4*	Axolotl *Dlx-3*
		E. coqui EcDlx3

Figure 1-4: Proposed model for the evolution of *dlx* genes.

An ancestral *distal-less* gene duplicated and inverted to generate a first pair of inverted convergent genes. Duplication of chromosome regions containing the primordial pair resulted in the existence of the six *Dlx* genes found in several vertebrate species. In zebrafish two supplementary *dlx* genes have been identified *dlx5* and *dlx8*. To date their chromosomal neighbours have not been identified


Ancestral
Distal-less
DlxA
DlxB
Dlx1
Dlx2
Dlx6
Dlx5 (dlx4)
Dlx7
Dlx3
dlx5
?
dlx8
?

1.4.3. Distal-less gene expression

1.4.3.1. *Drosophila*

The expression domains of *Dll* during *Drosophila* development seem to be restricted to the head and the appendages. Transcripts of *Dll* are first detected during cellularization of the blastoderm in a stripe that corresponds to the future maxillary and labial segments (Cohen, 1990). Prior to completion of cellularization, a second stripe corresponding to the antennal limb primordium appears. Following shortly after the onset of germ-band extension, expression is detected in the labrum primordium. By the completion of germ band elongation, the maxillary and labial primordia have expression domains resolved, the labrum primordium has migrated to the anterior tip of the embryo and the leg primordia have appeared and are also expressing *Dll*.

1.4.3.2. Vertebrates.

The expression pattern of the zebrafish *dlx* genes determined by *in situ* to whole mount embryos using antisens riboprobes to each *dlx* gene is summarised in table 1-2. The *dlx* genes seem to be involved during development of the forebrain (Porteus et al 1991; Price et al.,1991), olfactory placode (Akimenko et al 1994), branchial arches and craniofacial derivatives (Dollé et al., 1992; Papalopulu and Kinter, 1993; Akimenko et al., 1994), otic vesicle and inner ear (Ekker et al 1992a; Papalopulu and Kinter, 1993; Akimenko et al., 1994) and limbs/fins (Dollé et al.1992; Akimenko et al., 1994) (Figure 1-5).

One striking aspect of *dlx* expression is its combinatorial nature. It has been shown in zebrafish where *dlx1, dlx2, dlx4, dlx5* and *dlx6* are expressed in the ventral

forebrain with similar patterns. Similarly, *dlx3, dlx7, dlx4* and *dlx6* are expressed in the otic vesicle (table 1-2) (Akimenko et al., 1994; Ellies et al., 1997). There are extensive similarities between the overlapping expression of zebrafish *dlx* genes and that of their orthologs in other vertebrate species such as the mouse (Dollé et al., 1992; Porteus et al., 1991; Price et al., 1991) and *Xenopus* (Dirksen et al., 1993; Papalopulu and Kinter, 1993). This suggests that overlapping expression maybe an important aspect of Dlx function and is most likely the consequence of shared and conserved regulatory mechanisms that control multiple *dlx* genes.

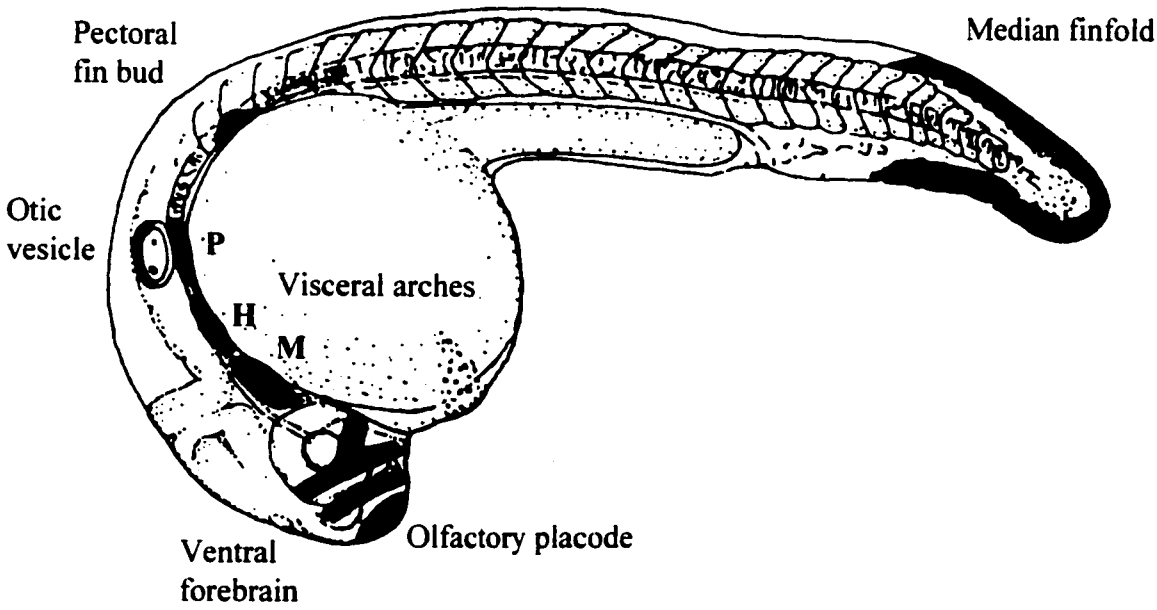
Partial or total genetic redundancy of genes that belong to a well conserved gene family has been proposed to explain the mild phenotypes or the absence of phenotype that have been observed in many selective gene inactivation experiments (gene knockouts) in the mouse. This is especially to be expected when expression patterns of these genes overlap such as is the case for the *Dlx* genes. The phenotype of *Dlx* null mutants in the mouse agree with this prediction. A null mutation in the mouse *Dlx2* gene was shown to result in abnormal morphogenesis of the proximal derivatives of the first and second branchial arches and abnormal differentiation in the forebrain, whereas a mutation that affects both *Dlx1* and *Dlx2* caused stronger phenotypes both in the arches and in the forebrain (Anderson et al., 1997; Qiu et al., 1997; Qiu et al., 1995). However, other structures such as the limbs, where the *Dlx2* and *Dlx1* genes are also normally expressed, are unaffected (Qiu et al., 1995). This suggests a partial genetic redundancy that may involve other *Dlx* genes, such as *Dlx3, Dlx5* and *Dlx6* which share overlapping expression patterns with *Dlx1* and *Dlx2* (Qiu et al., 1995). The abnormalities observed in *Dlx* null mutants suggest a unique function for each *Dlx* gene only in a subset of the cells where it is expressed.

Table 1-2: Summary of expression of the *dlx* genes during zebrafish development.

Ectodermal stripes in gastrula			*dlx3*	*dlx7*				
Ventral forebrain	*dlx1*	*dlx2*			*dlx4*	*dlx6*	*dlx5*	
Olfactory placodes			*dlx3*	*dlx7*	*dlx4*	*dlx6*		
Migrating neural crest		*dlx2*						
Visceral arches	*dlx1*	*dlx2*	*dlx3*	*dlx7*	*dlx4*	*dlx6*		*dlx8*
Dorsal otic vesicle			*dlx3*	*dlx7*	*dlx4*	*dlx6*		
Pectoral fin buds	*dlx1*	*dlx2*	*dlx3*	*dlx7*	*dlx4*	*dlx6*	*dlx5*	*dlx8*
Median fin fold	*dlx1*	*dlx2*	*dlx3*	*dlx7*	*dlx4*	*dlx6*	*dlx5*	*dlx8*

Figure 1-5: Domains of *dlx* expression in zebrafish embryo.

This is a sagittal view, dorsal to the top, anterior to the left, of a 24h zebrafish embryo showing the domains of *dlx* expression. At this stage, *dlx* genes are expressed in the dorsal otic vesicle, in the visceral arches, the olfactory placodes, the pectoral fin buds, the median fin fold and the ventral forebrain. In the ventral forebrain, two stripes of cells, which one of them shows a Y shape, correspond to the domain of expression. *Dlx* expression is also observed in the mandibular (M), hyoid (H) and pharyngeal (P) arches.


Pectoral
fin bud
Median finfold
Otic
vesicle
P
Visceral arches
H
M
Olfactory placode
Ventral
forebrain

1.4.4. *Dlx* homeobox genes as transcription factors: what is known thus far?

Little is currently known about the Dlx transcriptional properties vs Hox genes described before. It is thought that they act as transcriptional regulators since they have a homeodomain that can mediate DNA recognition. It is not clear to date what are the downstream targets of the proteins and how they interact with the responsive element of the target gene. In addition to genetic redundancy, it is suggested that one reason why *Dlx* expression patterns overlap is because Dlx proteins may cross-regulate each other as discussed previously. This hypothesis was verified by cotransfection experiments in cultured cells where several Dlx proteins were able to activate transcription of a CAT reporter containing an A/T-rich core sequence (Zerucha et al., 1997; Zerucha et al., 1999).

The functions of domains located outside the homeodomain of the family of *Dlx* genes is poorly understood. One such domain has been studied which is the activation domain. The transcriptional activation properties of a rat *Dlx3* were studied by Feledy et al (1999). They used cotransfection methods to identify the activation domains. They transfected HeLa cells with a mixture of two plasmids, one an expression construct in which the *Dlx3* coding sequence was driven by an RSV promoter and the other a reporter, the CAT coding sequence driven by a minimal promoter derived from the herpes simplex virus thymidine kinase gene (Wagner et al., 1981), in which three tandem *Dlx3* binding sites had been inserted 50 bp upstream from the start of transcription. In addition of transfecting full length *Dlx3*, they transfected deleted versions of it. They deleted the first 43, 93 and 130 amino acids and observed that an activation domain was located within amino acids 43 and 93 of the N-terminus region. They also deleted the last 16, 52 and 77 amino acids and observed an activation domain mainly located between amino acids 16 and 52 of the C-terminus. In summary, they found two activation domains in *Dlx3*, one in the N-terminal region and another one in the C-terminal region. They do not stipulate

though a possible molecular mechanism by which Dlx proteins can activate transcription (Feledy et al., 1999).

Dlx proteins are known to be activators of transcription but it was shown that they could act also as repressors. It has been reported that one of the rat Dlx members, Dlx5 was involved in the regulation of osteocalcin gene expression. Ryoo et al (1997) provides evidence to indicate that the rat Dlx5 functions to repress osteocalcin gene transcription. Their data provide an important demonstration that *Dlx5* expression correlates with osteoblast differentiation and suggest that *Dlx5* may be involved in maturation of the bone cell phenotype by acting as a repressor of transcription (Ryoo et al., 1997). This study is the only one to suggest that Dlx proteins act as repressors whereas numerous studies suggested that Dlx proteins act as activators of transcription.

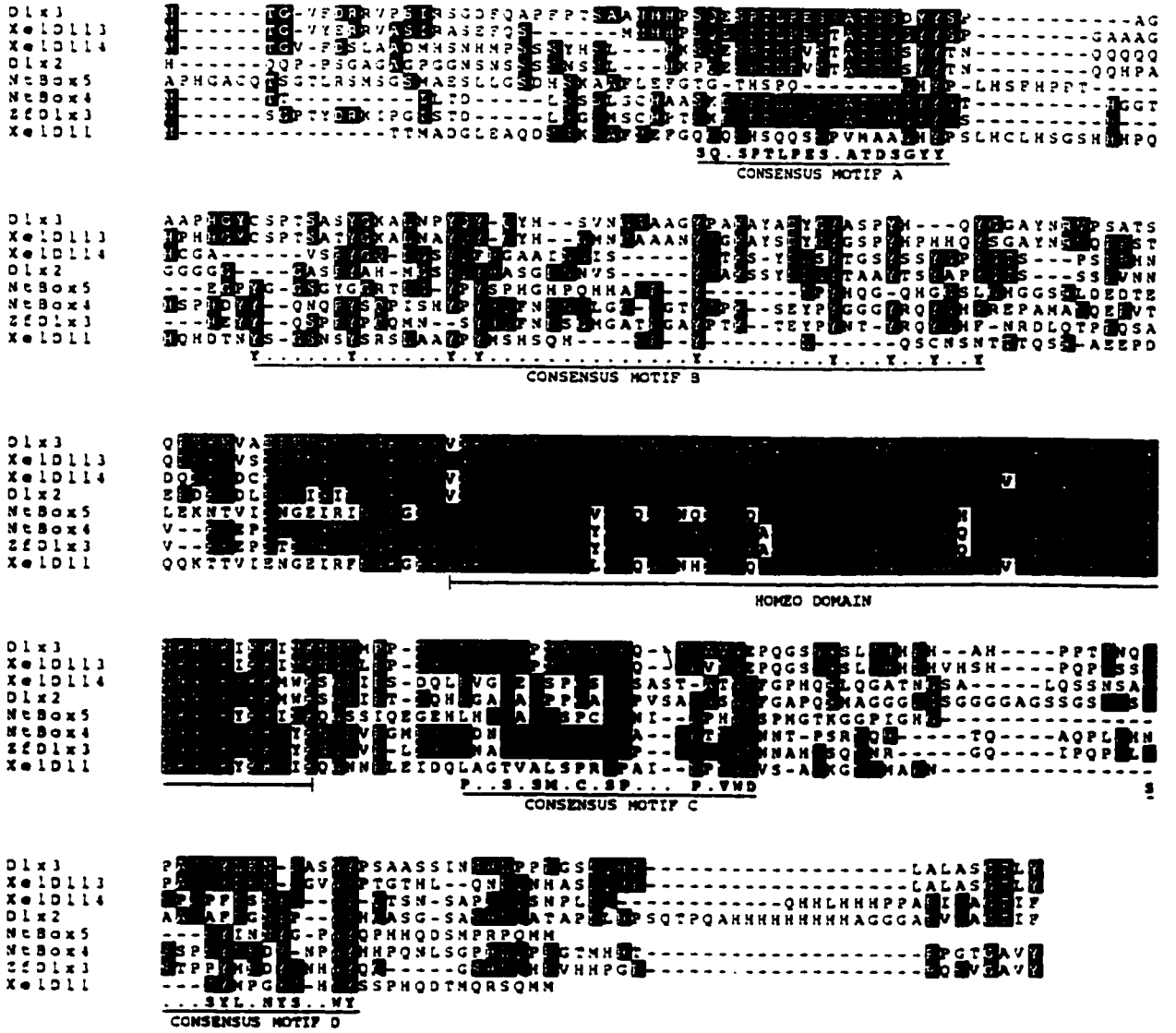
The molecular mechanism by which Dlx proteins exert their effect on target gene expression is poorly understood. It is suggested though that some Dlx proteins are able to form homodimers and heterodimers with some members of another class of homeodomain protein called Msx, which are expressed during embryogenesis and organogenesis. Zhang et al. (1997), demonstrated by GST interaction assays that Dlx (Dlx2 and Dlx5) and Msx (Msx1 and Msx2) can form homodimeric complexes between individual Msx and Dlx proteins (e.g: Msx1-Msx1), also that they can form heterodimeric complexes between members of the same family (e.g: Msx1-Msx2) and finally they can form heterodimeric complexes between members of the Msx and Dlx families (e.g: Msx1-Dlx2). They also showed that dimerization of Msx and Dlx protein is mediated through their homeodomains and that the residues required for this interaction correspond to those necessary for DNA binding. They demonstrated that the transcriptional properties of Msx and Dlx proteins display reciprocal inhibition. This means specifically that Msx proteins act as transcriptional repressors whereas Dlx proteins act as transcriptional

activators, while in combination, Msx and Dlx proteins counteract each other's transcriptional activities. This heterodimerization provides a mechanism for regulating the transcriptional activities of both Msx and Dlx homeoproteins *in vivo* (Zhang et al., 1997).

Characteristic motifs exist outside the homeodomain proteins, but it is not known if they are involved in transcriptional activation. Shirasawa et al studied the primary structure of Dll proteins and found an 18 amino acid residue stretch, designated as consensus motif A, which is found in the N-terminal region of Dll members (Figure 1-6). This motif is only present in Dlx homeoproteins but not in all families. Another motif (consensus motif B) is found in the N-terminal region of Dlx2, Dlx3, Dlx4 and Dlx5, which consists of 8-11 tyrosine residues that appear within 50 amino acids. Some of these tyrosine residues located at the interval of 7-8 amino acid residues or 14-15 amino acid residues, which would fit in 2 or 4 turns of an α helix, if this domain represents an α helix structure. This motif would be involved in protein-protein interactions. It has been suggested that the N-terminal region would be modulating (1) DNA-homeodomain interactions (Brand et al., 1993), (2) interact with effector molecules (Boulikas 1993), (3) interact with DNA (Perez et al., 1992), or (4) interact with other proteins to form the nuclear protein complex necessary for transcriptional activation (Gay et al 1988b). They also found 2 motifs, C and D, in the C-terminal region excluding the homeodomain. Those motifs are found to be serine rich residues which might be modified by protein phosphorylation Shirasawa et al., 1994). Further studies have to be done to clarify the function and role of Distal-less homeoprotein as a transcriptional activator.

Figure 1-6: Motifs found in the primary structure of a few Distal-less members

Amino acid alignment among members of Distal-less homeodomain protein. The amino acid sequence (in single letter-code) of rat *Dlx-3* has been aligned with those of *Xenopus Dll-3* (L09729, GenBank) (Shirasawa et al. 1994), newt *Box-4* (Beauchemin et al. 1992), zebrafish Dlx-3 (Ekker et al. 1992), *Xenopus Dll-4* (L09728, GenBank), mouse *Dlx-2* (Porteus et al. 1991), *Xenopus Dll* (Asano et al. 1992) and newt *Box-5* (Beauchemin et al. 1992) gene using a computer program contained in Lasergene; dashes denote gaps that have been introduced to maximize the alignment. Positions at which at least three of the sequences are identical are shown by black boxes.


Dlx1
Xelbl13
Xelbl14
Dlx2
NetBox5
NetBox4
zfDlx3
Xelbl1
SQ.SPTLPES.ATDSGYY
CONSENSUS MOTIF A
CONSENSUS MOTIF B
HOMEO DOMAIN
P..S.SM.C.SP... P.VWD
CONSENSUS MOTIF C
...SYL.NYS..WY
CONSENSUS MOTIF D

1.5. Aim of this investigation.

The *dlx* homeobox genes constitute a family of up to eight members in vertebrates that encode proteins that are highly similar in the DNA-binding homeodomain regions but very different outside the homeodomain. Based on the expression pattern, the *dlx* genes are thought to be involved during zebrafish development in the following regions of the embryos: otic vesicle, olfactory placode, fins, forebrain and the visceral arches. Previous work done in the laboratory (zebrafish *dlx* genes) and other laboratories (*Dlx* genes from vertebrates other than zebrafish) has shown that multiple *dlx* genes are often co-expressed in the same cells during development. To better understand the functional specificity and the biochemical properties of Dlx proteins, we decided to use transgenic zebrafish embryos and experiments in cultured cell systems to identify the regions of Dlx proteins that are important for their function and determine how the various Dlx proteins differ from each other. Transfection in cultured cells have been done to identify Dlx domains responsible for transcriptional activation. I demonstrated in this thesis that an activation domain was located in the N-terminal region of Dlx1, Dlx2 and Dlx4. In addition, mRNA injections of truncated *dlx3* and a chimera of *dlx3-dlx2* have been done in zebrafish embryos at one-cell stage. I showed that a chimera version constituted of the N-terminal half of Dlx3 and the homeodomain and C-terminal half of Dlx2, was perturbing endogenous *dlx4* expression and subsequently resulted in malformations of the craniofacial cartilages.

2- Material and Methods

2.1. Subcloning dlx genes in expression vector.

The *dlx* genes chosen to study the activation domain were *dlx1*, *dlx2* and *dlx4*. In order to locate those activation domains in these genes I dissected them, so for each gene the regions upstream and downstream of the homebox would be individually investigated. Some of these fragments were subcloned from a yeast vector PasII (figure 2-1) to an expression vector PTL1 (figure 2-2). It was easier to do so, since they were already fused in frame to a Gal-4 DBD. *Dlx* cDNA's had been previously subcloned in the yeast vector by another member of the laboratory in order to study the activation domains in Dlx proteins using the yeast one hybrid system. We decided to study the activation domains in Dlx in mammalian cultured cells instead of yeast so it would be more representative of what occurs in vertebrate and also allowing us to obtain quantified results. Some *dlx* fragments that were not subcloned previously in PasII were subcloned into the expression vector GalO which contains the same Gal-4 DBD sequence as the one found in the PasII vector.

Dlx1 full length

Dlx1 fused in frame downstream of a Gal-4 DBD was cut from the PasII vector using the *Hind*III and *Sal*I sites. The fragment was then ligated in the PTL1 (figure 2-2) mammalian expression vector using the *Hind*III and *Xho*I site. Since no *Sal*I site was present in the PTL1 vector, we used *Xho*I site because the digested fragment ends produced by *Sal*I and *Xho*I are compatible.

Figure 2-1: Map of the expression vector PasII

This vector contains a Gal-4 DNA binding domain (1-147) just before the multiple cloning site. Some *dlx* fragments (full length *dlx1*, *dlx1* N-terminal, *dlx1* N-terminal*, *dlx2* N- and C-terminal, *dlx4* and *dlx4* C-terminal) were subcloned from PasII to a mammalian expression vector PTL1.

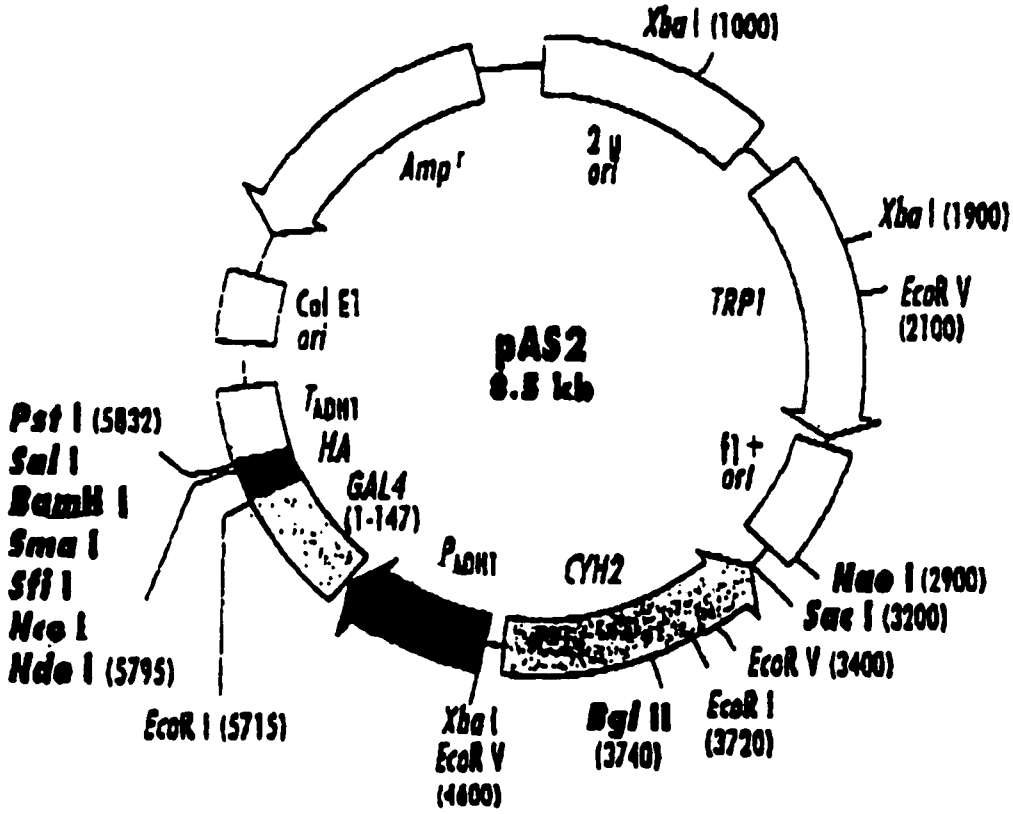
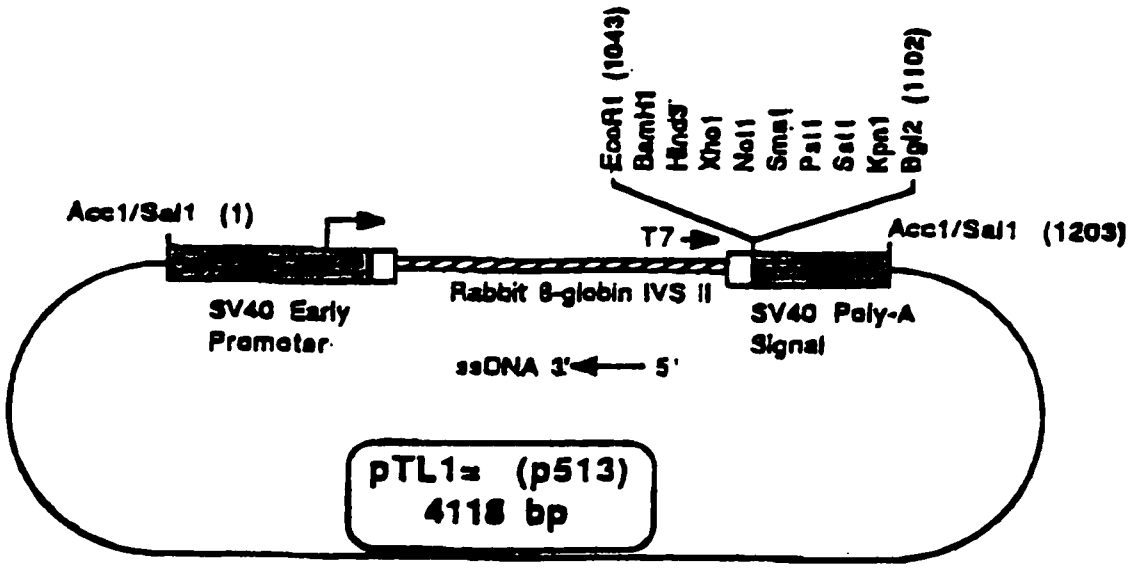

Xba I (1000)
Amp r
2 μ ori
Xba I (1900)
EcoR V (2100)
Col E1 ori
TRP1
pAS2
8.5 kb
Pst I (5832)
Sal I
BamH I
Sma I
Sfi I
Nco I
Nde I (5795)
T ADH1
HA
GAL4 (1-147)
f1 + ori
Nae I (2900)
Sac I (3200)
P ADH1
CYH2
EcoR V (3400)
EcoR I (5715)
Xba I
EcoR V
(4600)
Bgl II
(3740)
EcoR I
(3720)

Figure 2-2: Map of the expression vector PTL1

This vector contains an SV40 early promoter which can allow transcription of genes inserted in the multiple cloning site in mammelian cell system. The following *dlx* genes and fragments were subcloned in PTL1 along with the Gal-4 DBD present in PasII: full length dlx1, dlx1 N-terminal, dlx1 N-terminal*, dlx2 N- and C-terminal, dlx4 and dlx4 C-terminal.

Expression vector PTL1



Dlx1 amino-terminus

Then, the *dlx1* 5' part of the gene which correspond to the first 417 base pairs of *dlx1* cDNA was cloned in the same way as full length *dlx1*. I cut the gene fusion of Gal-4 and full length *dlx1* with *Hind*III and *Pst*1 giving a 959 base pair fragment. The *Pst*1 site is located in the homeobox, 37 base pairs after its beginning. This fragment was then ligated in the expression vector using the same sites.

Dlx1 shorter amino-terminus*

Another *dlx1*-5' fragment shorter than the previous *dlx1*-5' fragment was also subcloned. The shorter *dlx1*-5' fragment was amplified by PCR using as template the yeast Gal-4 DBD fused upstream of full lenght *dlx1*. To achieve this, I used again primers containing appropriate restriction sites like *Hind*III and *Bgl*II. This fragment was then ligated in the digested PTL1 vector.

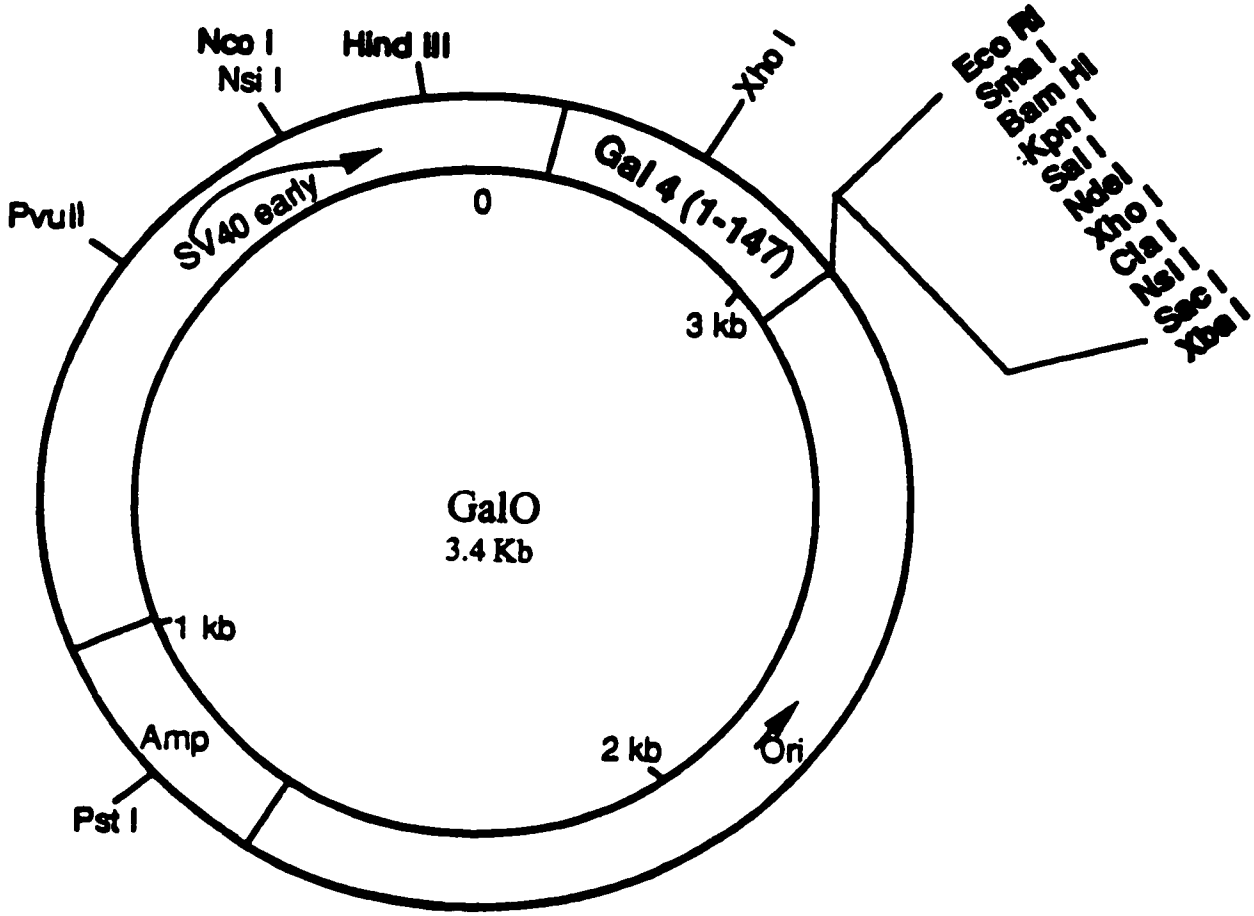
Dlx1 carboxy-terminus

The 3' fragment which starts just before the homeobox of *dlx1* was also subcloned. This time it was not possible to subclone it directly from the yeast vector. I had to use the PCR method to amplify the fragment using primers with endonuclease restriction sites inserted at their extremities. The restriction site sequences added to the primers were *EcoR*I and *BamH*I. The amplified fragment was then ligated in frame into the GalO expression vector, which already contains a Gal-4 DBD (Figure 2-3). The Gal-4 DBD of this vector is the same as the one found in the yeast PasII vector (amino acid positions 1-147). The length of the Gal-4 and *dlx1*-3' fragment is 913 base pairs.

Figure 2-3: Map of the expression vector GalO

GalO contains a Gal-4 DNA binding domain (1-147) located just before the multiple cloning site. The transcription is under the control of an SV40 early promoter. The *dlx1* C-terminus, full length *dlx2* and its longer C-terminus as well as *dlx4* N-terminus were subcloned in frame with the Gal-4 DBD of this vector (see table 2-1 and 2-2).

Expression vector GalO



The other *dlx* genes and fragments were similarly subcloned see table 2-1 and table 2-2 for more details of how each fragments were subcloned.

2.2. Co-transfection of *dlx* genes and CAT reporter.

The expression vectors were co-transfected with a reporter construct G5E1b CAT containing 5 Gal-4 binding sites and a TATA box and the chloramphenicol acetyl-transferase (CAT) gene (figure 2-4 and 2-5). For the transfection experiments, I used the Chinese Hamster Ovary Cell line and lipofection as a transfection procedure. There is at present no evidence that CHO cells express *dlx* genes endogenously. The cells were plated 24 h prior to transfection at a density of $2.5X10^5$ cells per 60mm dish. In each experiment I transfected 1μg of the expression vector, 0.5μg of the reporter CAT and 100ng of the internal RSV*β*Gal control plasmid to normalize for transfection efficiency. Ten microliters of GIBCO lipofectamine were added for each transfection. These were taking place in 400μl of serum free Minimum Essential Medium Alpha media. After 30 minutes of incubation, the mix was poured on confluent cells in 60mm tissue culture dishes. The DNA-lipofectamine complex was left on the cells for 5 hours. At this time, media containing 10% fetal bovine serum was added to the cells to neutralize the lipofectamine and avoid toxicity. The cells were harvested in 1X PBS 48hrs post-transfection, and pelleted by centrifugation (1min. at 7000rpm). The cells were then resuspended in 150μl of a freeze-thaw buffer (250 mM Tris-HCl, pH 8, 10 mM dithiothreitol, 15% glycerol). Whole cell extracts were prepared by repeated freezing and thawing.

Table 2-1: The *dlx* cloning strategies

This table lists all the restriction enzymes used for cloning *dlx1*, *dlx2* and *dlx4* full length as well as their respective N- and C-terminal regions. Also the amino acid position where these enzymes cut is indicated. The names of the oligonucleotides used for cloning *dlx1* C-terminus, *dlx1* N-terminus*, full length *dlx2* and its C-terminus*, and *dlx4* N-terminus in GalO are also listed. For more details about the oligonucleotide sequences see table 2-2. The GenBank accession numbers for *dlx1*, *dlx2* and *dlx4* are respectively U67842, U03875 and U03876.

Construction	Endogenous restriction site used (Position in aa sequence deposited in Genbank)		Oligonucleotide used
dlx1	*Hind*III (in PasII)	1	
	*Sal*I (in PasII)	251	
dlx1 N-term.	*Hind*III (in PasII)	1	
	*Pst*I (in dlx1 cDNA)	139	
dlx1 C-term.	*EcoR*I (dlx1a)	121	dlx1a
	*BamH*I (dlx1b)	251	dlx1b
dlx1 N-term.*	*Hind*III (in PasII)	1	dlx1c
	*Bgl*II (in dlx1d)	105	dlx1d
dlx2	*BamH*I (in dlx2a)	1	dlx2a
	*Bgl*II (in bluescript)	270	dlx2b
dlx2 N-term.	*Hind*III (in PasII)	1	
	*Pst*I (in dlx2 cDNA)	101	
dlx2 C-term.	*Hind*III (in PasII)	123	
	*Sal*I (in PasII)	270	
dlx2 C-term*	*BamH*I (in dlx2c)	96	dlx2c
	*Bgl*II (in bluescript)	270	dlx2d
dlx4	*Hind*III (in PasII)	1	
	*Sal*I (in PasII)	283	
dlx4 N-term.	*BamH*I (in dlx4a)	1	dlx4a
	*Bgl*II (in dlx4b)	107	dlx4b
dlx4 C-term.	*Hind*III (in PasII)	77	
	*Pst*I (in PasII)	283	

Table 2-2: Oligonucleotides used for cloning *dlx1*, *dlx2* and *dlx4* fragments.

All the sequences in this table start at the 5' extremity and end at the 3' extremity. Also, the restriction enzymes used for cloning are identified.

Oligonucleotide	**Sequence**
dlx1a	CG GAA TTC AAC GGG AAA GGC AAA AAG *EcoR*I
dlx1b	GC CCT AGG AGC AGT CCT GGT TCA CAT GAG *BamH*I
dlx1c	CG GGA TCC CCA AGC TTG AAG CAA GCC *BamH*I
dlx1d	CT TCT AGA TGC TGG GTC CTC CAA TCG *Bgl*II
dlx2a	CG GGA TCC GC GGT ATG AAA AAC ATG ACT *BamH*I
dlx2b	GAT ATC ACT CAG CAT TAA T7
dlx2c	CGC GGA TCC CC TGC TCC TCA CCA ACT *BamH*I
dlx2d	GAT ATC ACT CAG CAT TAA T7
dlx4a	CG GGA TCC GT CTT ATC CGA ACT ATG ACT *BamH*I
dlx4b	CT TCT AGA TGT TCC TGC GTA TTG GTG *Bgl*II

2.3. CAT assay and *b* Gal assay.

For the CAT assay, each sample was composed of 70 µl of the cell extract, 1µl of ^{14}C chloramphenicol, 20 µl of actetyl coenzyme A and 28 µl of 0.25 M Tris-HCl, pH 7.8. The reaction mixture was incubated at 37°C for 4 hours. The reaction was stopped by adding 800µl of ethyl acetate. The samples were then mixed and centrifuged at maximum speed (1400 RPM) for 5 minutes to separate the phases. The upper phase (700µl) was transfered to a fresh eppendorf tube and evaporated to dryness in a Speedvac (about 15 minutes). Finally, the "precipitate" was resuspended in 8µl of ethyl acetate and spotted onto a TLC plate. It was developed in a tank containing 100ml of a freshly made mixture of chloroform:methanol (95:5).

The *B* Gal assays are used as an indicator of transfection efficiency. For each sample, approximately a third of the cell extract (40 µl) was used to do the assay. To this, 900 µl of a solution made by combining 10 ml of sodium phosphate (60mM Na_2HPO_4, 40mM Na H_2PO_4), 100µl of 1M potassium chloride, 10µl of 1M magnesium chloride and 35µl of *B*-mercaptoethanol were added. To start the reaction 200µl of ONPG (o-Nitrophenyl *B*-D-Galacto-pyranoside, 4 mg/ml in sodium phosphate solution described above) were added to each sample. ONPG is a substrate for the *B*Gal protein. The samples were mixed and incubated at 37°C and the time zero of reaction was noted. The reaction was stopped at various times afterwards with 500µl of sodium carbonate 1M. The reaction product was measured spectrophotometrically (O.D. at 420 nm).

The numbers were calculated from the analysis of ^{14}C-chloramphenicol counts and the converted ^{14}C-chloramphenicol were determined by the software Molecular Analyst. The CAT activity numbers were determined as followed:

Figure 2-4: Reporter plamid G5 E1B CAT map

The G5 E1B CAT reporter plasmid was used for transfection experiments. This plasmid contains 5 Gal-4 DNA binding sites, a TATA box and the CAT gene. Nothing was cloned in the multiple cloning site.

Reporter plasmid

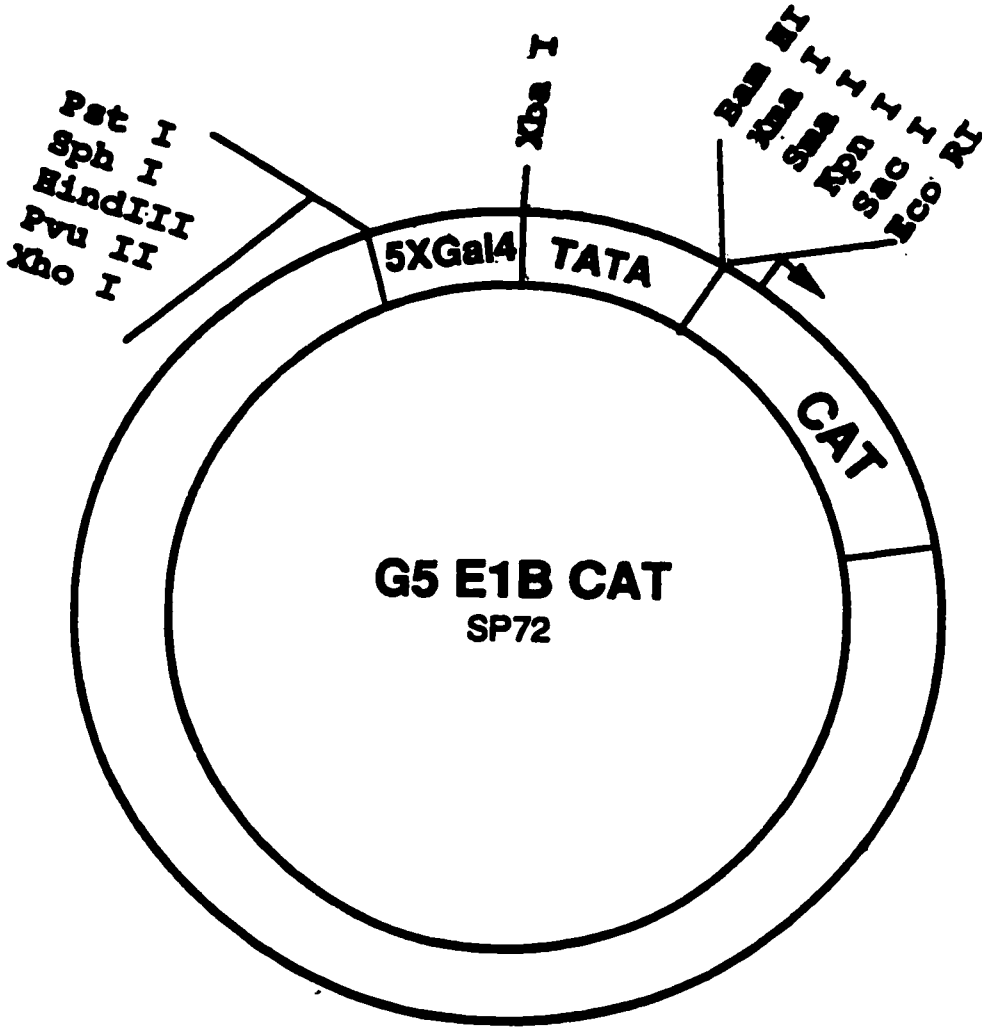
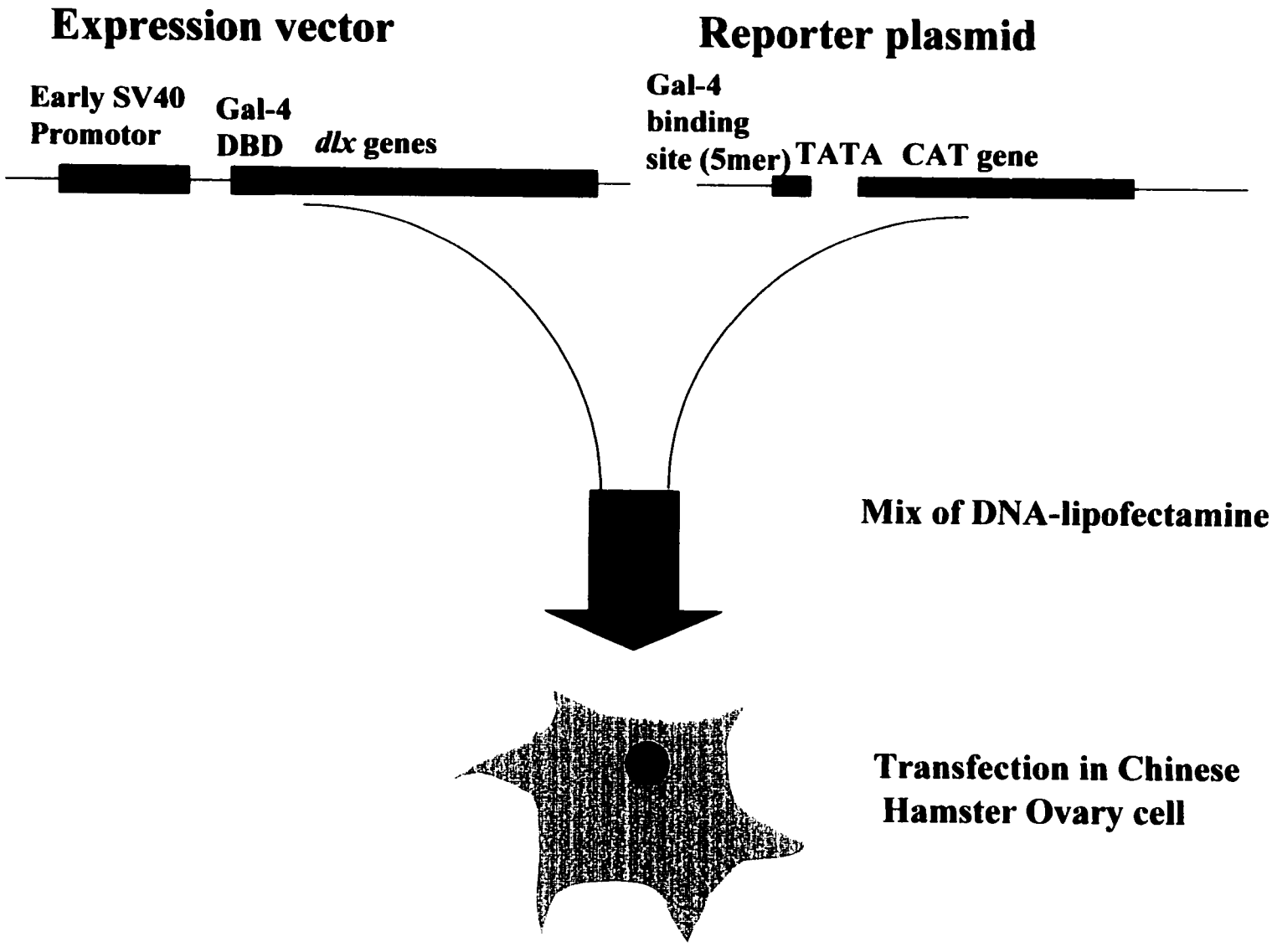


Figure 2-5: Transfection procedure.

The *dlx* genes or their fragments were subcloned in an expression vector (PTL1 or GalO) fused in frame to a Gal-4 DBD. This vector was co-transfected along with the reporter plasmid G5 E1B CAT. The protein products of the Gal-4 DBD fused to the Dlx protein or protein fragments can bind to the 5 Gal-4 DNA binding sites of the reporter. The presence of activation domains in the full length Dlx protein or Dlx protein fragments will allow transcription of the CAT gene.


Expression vector
Reporter plasmid
Early SV40
Promotor
Gal-4
DBD
dlx genes
Gal-4
binding
site (5mer)
TATA
CAT gene
Mix of DNA-lipofectamine
Transfection in Chinese
Hamster Ovary cell

$$\text{CAT activity} = \frac{\%\,^{14}\text{C-chloramphenicol conversion}}{\text{Conversion factor}}$$

$$\%\text{Conversion} = \frac{^{14}\text{C-chloramphenicol converted counts}}{^{14}\text{C-chloramphenicol total counts}} \times 100$$

$$\text{Conversion factor} = \frac{\text{x } \textit{B}\text{gal factors}}{\text{Lowest } \textit{B}\text{gal factors}}$$

$$\textit{B}\text{gal factor} = \frac{\text{(O.D. 420/ 0.0045)}}{\text{Time}}$$

2.4. Western analysis

2.4.1. Micropreparation for extraction of DNA-binding proteins.

The cells were plated 24 h prior to transfection at a density of $2.5X10^5$ cells per 60mm dish. The transfection was done as described above, with the exception that only the expression vector was transfected, at a concentration twice higher than that used in the experiments involving CAT assays. After 24 h, the cells were harvested in 1ml PBS and centrifuged for 1 minute at 7000 RPM. The supernatant was removed and the cells resuspended in 400µl of cold buffer A (10mM Hepes-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5mM DTT and 0.2 mM PMSF) by flicking the tube. Cells were left to swell on ice 10 minutes, and then mixed for 10 seconds. The extract was then centrifuged 10 seconds at 440 rpm and the supernatant was discarded. After the pellet was resuspended in 50 µl of cold buffer C (20 mM Hepes-KOH pH 7.9, 1.5 mM $MgCl_2$, 420mM EDTA, 0.5 mM DTT, 0.2 mM PMSF and 25% glycerol), it was incubated on ice for 20 minutes.

The cellular debris was centrifuged for 2 minutes at 4°C at maximum speed in a microcentrifuge. Finally the samples were stored at -80°C.

2.4.2. Western Blot

Half of all samples (20µl) were electrophoresed by SDS-PAGE on a 15% gel. Loading of equal amounts (µg) of proteins in each well would have been preferable. However, protein concentrations were not determined. It must be mentioned that all plates of cells in the experiments had grown to comparable levels of confluence. Therefore, it is relatively safe to assume that equal weights of protein were loaded on the gel. The proteins were then transferred to a PVDF membrane using transblot buffer (25mM Tris-HCl pH 7.7 at room temperature, 192 mM glycine, 20% methanol and 0.1% SDS). After the transfer, the PVDF membranes were rinsed in TBS-tween buffer (20mM Tris, pH 7.5, 500 mM NaCl, 0.5% tween). Then, the membranes were blocked with 10% skim milk TBS-tween (2.5g in 25ml), for 1 h at room temperature with gentle mixing. The membranes were washed 15 minutes at room temperature with TBS-tween and incubated overnight at 4°C, with the primary mouse monoclonal Gal-4 DBD IgG_{2a} antibody (Gal-4 DBD antibody) (1:2000 dilution in TBS-Tween), in a sealed bag. The next day the membranes were washed 3 times at room temperature, (once for 15 min, twice for 5 min). The secondary antibody, sheep antimouse conjugated to horse radish peroxidase (HRP), 1:50 000 dilution in TBS-tween, was added and the membranes incubated for 1 h at room temperature with gentle mixing. This was followed by 2 washes, 2h each, in TBS-tween buffer. The membranes were then soaked in an ECL mixture (1:1) for a minute, and finally, the film was exposed to membranes overnight.

2.5. Animals

Wildtype zebrafish embryos were bought from a pet store and were maintained under conditions conforming to standard procedures (Westerfield,., 1993). Embryos were staged at 28.5°C according to hours (hpf) and days (dpf) post fertilization.

2.6. Subcloning dlx genes in CS2+

The *dlx2* and *dlx3* full length cDNAs were subcloned from bluescript using convenient restriction sites into the CS2+ vector (figure 2-6). The same procedure was done for the subcloning of the truncated forms of *dlx2* and *dlx3* and for the chimera. Those fragments were first cloned in bluescript, but we decided to subclone them in the CS2+ vector containing a SV40 poly A tail that confers more stability to the mRNA synthesized (figure 2-6). A truncated *dlx3* was also created using a *Xho*I site in the 5' region of the *dlx3* homeobox and codes for a 131 amino acid polypeptide, including the amino-terminus and the first nine amino acids of the homeodomain. The truncated *dlx2* was subcloned in the CS2+ vector from the *dlx2* cDNA in bluescript using the *Xho*I site found at the same position in the homeodomain as for *dlx3*. The *dlx3-dlx2* chimera, was created by making use of this conserved *Xho*I site in the homeoboxes of *dlx2* and *dlx3*. The creation of the chimera does not affect the reading frame of *dlx2* nor *dlx3*. The chimeric protein consists of the amino-terminus and the first nine amino acids of the homeodomain of Dlx3; the remainder of the homeodomain and carboxy-terminus are coded by Dlx2 sequences.

2.7. mRNA synthesis

The templates were linearized using *Asp*718 (figure 2-6). In vitro synthesis was done using the mCAP mRNA capping kit (Stratagene). The viral Sp6 RNA polymerase was used for transcription (see figure 2-6). Synthetic RNA was quantified by absorbance at 260nm.

2.8. Injection of mRNA into developing zebrafish.

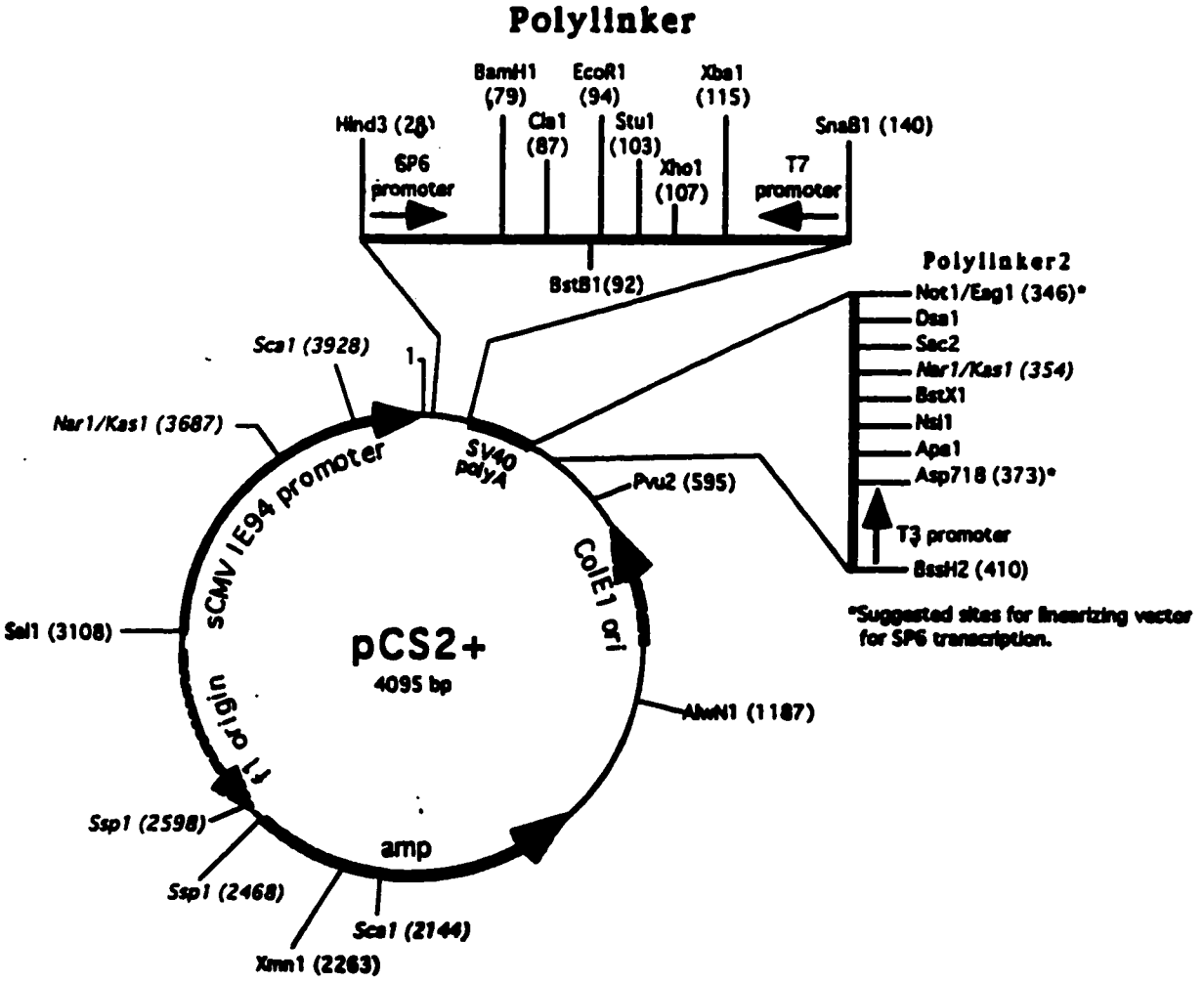
Synthetic capped RNA was injected into zebrafish embryos at the one-cell stage using the procedures described in Stuart et al. (1988). Briefly, synthetic capped RNA was dissolved in distilled water containing 0.1% phenol red to visualize and the volume of RNA solution injected approximately was estimated to be 10 nl. Levels of expression of mRNA were not monitered. A control consisting of a synthetic mRNA labelled with digoxigenin was previously injected and found to be ubiquitously distributed, 24 hours after injection.

Embryos were kept for 27 hours in embryo medium. To prepare embryo medium add: 1ml of Hank's solution 1 (8g NaCl, 0.4g KCl im 100ml ddH_2O), 0.1ml of Hank's solution 2 (0.358g Na_2HPO_4 Anhydrous, 0.60g KH_2PO_4 in 100ml of ddH_2O), 1ml of Hank's stock solution 4 (0.72g $CaCl_2$ in 50ml H_2O), 95.9ml ddH_2O, 1ml of Hank's solution 5 (1.23g $MgSO_4$ $7H_2O$ in 50ml ddH_2O), 1ml of fresh Hank's solution 6 (0.35g $NaHCo_3$ in 10.0ml ddH_2O), finally, use about 10 drops 1M NaOH to pH 7.2.

Figure 2-6: Map of the expression vector CS2+

The complete *dlx3* and *dlx2* cDNA's as well as a chimera consisting in the amino-terminal half of *dlx3* and the homeodomain and carboxy-terminal half of *dlx2*, truncated *dlx2* and *dlx3* were subcloned in the first of the two polylinkers using convenient restriction sites. The transcription is under the control of the sCMV IE94 promoter. A SV40 poly A site is located downstream of the polylinker where cDNAs were inserted. This is thought to confers more stability to the synthesized mRNA.

Expression vector CS2+



2.9. Whole mount *in situ* hybridization

After 27 hours of development, the expression of endogenous *dlx* genes in the injected embryos was studied using the *in situ* hybridization technique. The antisense RNA digoxigenin (DIG)-11-UTP (Boehringer Mannheim) probe was synthesized from a linearized cDNA template. Briefly, the linearized DNA (1μg) template is incubated for 2 hours at 37°C with NTP's (0.073 mmol), DIG-11-UTP (7 nmol), and either T3 or T7 RNA polymerase (20 units). The DNA template is then removed from the mixture by adding 1 ul DNase. The DIG-labelled fragment is then precipitated with ethanol/LiCl overnight at -20°C. The precipitated probe is partially hydrolysed, to facilitate probe penetration, with sodium bicarbonate/sodium carbonate at 60°C. The time of hydrolysis depends on the length of the labelled fragment.

$$\text{Time of hydrolysis} = \frac{\text{start length - desired length}}{0.11 \text{ X (start length X final length)}}$$

The following templates were used for synthesizing the antisense riboprobes: *dlx2*, a cDNA fragment of 1667 base pair, *dlx3*, a 1532 base pair fragment and *dlx4*, a 1123 base pair fragment (Akimenko et al. 1994).

Zebrafish embryos were fixed in 4ml of 4% paraformaldehyde (PFA) for 2 hours at room temperature. *In situ* hybridizations were performed on whole mount embryos using antisense riboprobes as described above. All steps below are performed using RNase free solutions and at room temperature unless otherwise indicated. The fixed embryos are rehydrated in serial baths of 75%/25%, 50%/50% and 25%/75% methanol/RNase-free water. Once rehydrated, 1/6 of the embryos were used for preabsorption of the anti-DIG antibody. The embryos are transferred in a 1.5ml eppendorf

with the anti-DIG antibody conjugated to alkaline phosphatase diluted 1/1000 in 940 µl of PBST, 2% calf serum (20 µl), 2mg/ml bovine serum albumine (20µl). The preabsorption step is carried out at room temperature for 2 hours then overnight at 4°C.

The *in situ* embryos are permeabilized with proteinase K (4 µl of 20ug/ml in 4ml PBS-tween (1X PBS, 0.1% tween-20)) (PBST) for 2 minutes. The embryos are then postfixed for 20 minutes in 4ml of 4% PFA, and rinsed twice in 4ml PBST. To minimize background, an acetylation step is performed for 10 minutes. The acetylation mixture is made of 125µl triethanolamine, 27µl acetic anhydride, in 10ml RNase-free (RF) water. The embryos are rinsed in 4ml of PBST for 10 minutes twice.

After the embryos are incubated 2 hours at 65°C in a prehybridization buffer (50% deionized formamide (5ml), 5XSSC (2.5 ml of 20X SSC), 0.1% tween-20 (50 µl of 20%tween-20), 50 µg/ml heparin (100 µl of 5 mg/ml), citric acid (92 µl of 1M stock pH 6.0), add water to 10 ml and 100 µg/ml yeast tRNA), they are hybridized, overnight, in the same solution supplemented with 100 ng (0.5 ng/ul) of labelled RNA probe. The next day the hybridized embryos are washed at 65°C for 10 minutes in sequential baths containing 4 ml solution composed of decreasing percentage of hybridization solution and increasing percentages of 2X SSC, starting with 75% of the hybridization mix without probe (3 ml)+ 25% 2X SSC (1 ml), then 50% of both (2 ml each) and 25% of hybridization mix (1 ml) and 75% of 2X SSC (3 ml), until the embryos are immersed in 4 ml 2X SSC solution. They are washed twice for 30 minutes each in 4 ml 0.2X SSC at 60°C. The embryos are brought to room temperature and put through sequential baths containing decreasing percentages of 0.2X SSC and increasing percentages of PBST starting with 75% SSC (3 ml)+ 25% PBST (1ml) and so on, until they are transferred in 4 ml PBST.

The hybridized embryos are preincubated for 1 to 4 hours in 4ml PBST, 2% calf serum (80µl), and 2mg/ml bovine serum albumine (80µl). After the preincubation, the embryos are incubated for 3 to 4 hours with 1/5000 dilution of the preabsorbed antibody (800µl of preabsorbed antibody solution in 4.2 ml of PBST). This reaction allows the anti-DIG antibody to bind to the DIG labelled probe. The embryos are washed in 4 ml PBST, and again preincubated in a 4 ml prestain buffer (100mM Tris-HCl pH 9.5 (1.5 ml); 50mM $MgCl_2$ (750µl); 100mM NaCl (300µl); 0.1% Tween-20 (75µl of tween-20); 1mM levamisol (150µl of a 100mM stock)) for 5 minutes. The embryos are then stained using the substrates 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) (14µl of 50mg/ml stock solution) and 4-nitrophenyl phosphate (NBT) (27µl of 50mg/ml stock solution). A chromogenic precipitate forms when the stain reacts with the alkaline phosphatase conjugated to the anti-DIG antibody, leaving a blue-purple color precipitate. A 2 hours post fixation in 4 ml 4% PFA followed by a wash with 4 ml of PBS is carried out. The embryos are then stored at 4°C in 100µl of PBS/5mM sodium azide.

2.10. Alcian blue staining

Alcian blue is a substance which is able to stain the cartilage. Briefly, 5-days larvaes were fixed in 4 ml 4% PFA and washed in 4 ml PBS. They were stained for 5 hours in 4 ml of a stock solution composed of 15mg alcian blue dissolved in 7 ml of ethanol and 3 ml of acetic acid. The embryos were then rehydrated in a 5%water-95% ethanol solution. Stained larvae were kept in glycerol at 4°C.

2.11. Phalloidin staining

The five-day larvae fish were fixed in 4ml of PFA for 2 hours at room temperature. They were rinsed in 4ml PBS and soaked in a PBS-Triton (2%) solution for

24 hours. The larvae were then stained in phalloidin conjugated to alexa^{TM} 488 (Molecular Probes) (5µl in 200µl PBS) at room temperature for 2 hours. They were washed in 4 ml PBS for 2 hours. The dish containing the larvae were kept in aluminum foil at all stages to protect them from light. Just before the visualization on the confocal microscope (excitation at 522 nm), the larvae were stained in ethidium bromide/PBS solution (dilution 1/200).

3- Results:

3.1. Activation levels of Dlx1, Dlx2, Dlx4 and their respective fragments.

In order to study the activation domains of different Dlx's, a co-transfection experimental strategy was used. *dlx* cDNA or cDNA fragments were subcloned in an expression vector (PTL1 or GalO), downstream and in phase of a Gal-4 DBD. Transcription factors are often modular in nature. Therefore, it is possible to fuse functional domains (DBD, activation domains) from different proteins into a single protein. The relative order of the different functional domains will not influence function in many of the reported examples. These constructs are co-transfected in CHO cells with a CAT reporter construct containing 5 Gal-4 binding sites, a TATA box and the CAT gene. In the cells the Gal-4 DBD-Dlx fusion protein is produced and can bind to the Gal-4 binding sites through the action of the Gal-4 DBD. If the Dlx fragment contains an activation domain, the CAT gene is going to be transcribed at levels related to the strength of the activation domain.

The relative CAT activities conferred by Gal-4 DBD-Dlx fusion proteins containing the complete coding regions of *dlx1*, *dlx2* and *dlx4* as well as various fragments of the above proteins were compared. *Dlx1*, *dlx2* and *dlx4* genes were chosen for this study in order to have members of each of the two main clades (see figure 1-4). It was reasonable to predict that the main differences in activation properties would be found between these two clades. In addition, these genes were the most studied and accessible in the laboratory. The numbers were calculated on the basis that full length Dlx1 activation is equal to 100%. The activities of the two controls which consisted in transfecting the expression vector PTL1 alone or transfecting the expression vector PTL1 containing the Gal-4 DBD only, were calculated based on the activation of Dlx1 and

equal to 6% and 7% respectively. The N-terminus of Dlx1 is a protein fragment of 139 amino acids. This fragment spans from the N-terminus to the 12^{th} amino acid of the homeodomain (figure 3-1). This fragment produces CAT activities 2.5 times higher than the full length Dlx1fused to Gal-4 DBD. In contrast the C-terminus of Dlx1, which spans from the homeodomain to the stop codon (130 amino acids) activated the transcription at lower level than Dlx1 N-terminus which gives 18% of the CAT activity full length Dlx1. This fragment activates transcription 2.5 times higher than the controls (figure 3-2, table 3-1)

We also tried to identify the activation domain of Dlx2. Full length Dlx2 which is a protein of 270 amino acids activates transcription of the reporter to levels comparable to Dlx1 (98%) (figure 3-3 and table 3-1). The N-terminal fragment of Dlx2 contains 101 amino acids, starting at amino acid 1 and ending at 22 amino acids before the homeodomain (amino acid position 123) (figure 3-1). This fragment activates transcription of the reporter at levels 41% those of full length Dlx1 (figure 3-3 and table 3-1). This fragment is not activating as much the transcription as the full length Dlx2, which contrasts with the behavior of the N-terminus of Dlx1. The C-terminal fragment contains the complete homeodomain (amino acid position 123) to the C-terminus of Dlx2 (270) (figure 3-1). The activation of the C-terminus is 22% of Dlx1 and is 3 times higher than the activity of the controls (figure 3-3 and table 3-1).

Another member of the Dlx family, Dlx4, was also studied for its activation domain. The full protein is 283 amino acids long and its activity is low compared to Dlx1 and Dlx2, 29% (figure 3-3 and table 3-1). The N-terminus and C-terminus of Dlx4 were also studied. The N-terminus of Dlx4 spans the first 107 amino acids, and does not contain the homeodomain (figure 3-1).

Figure 3-1: *Dlx* fragments used for transfection assays

The first amino acid and the last amino acid positions are indicated for each constructs. The position of the homeodomain is represented by the gray rectangle.

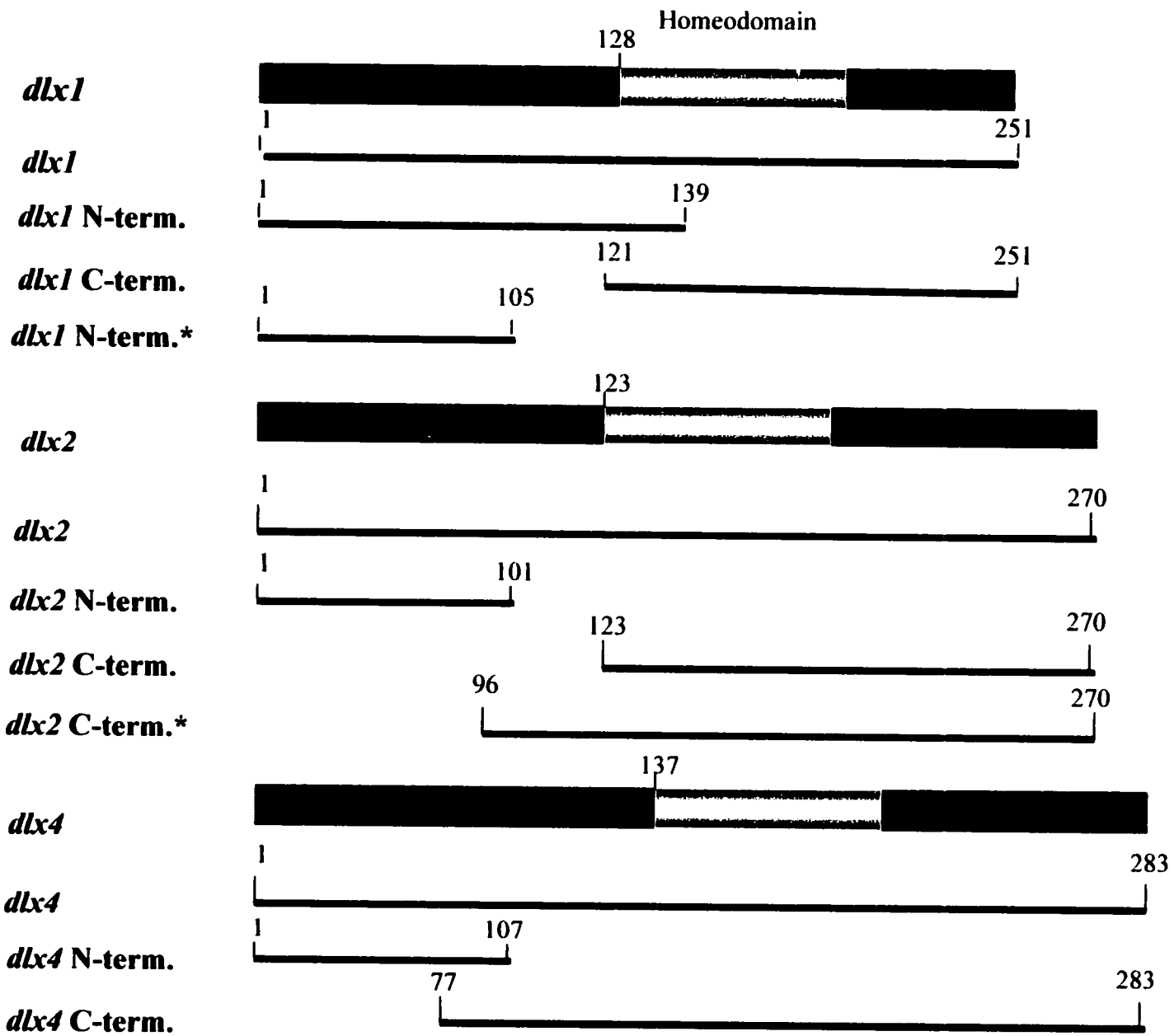

Homeodomain
128
dlx1
1
251
dlx1
1
139
dlx1 N-term.
121
251
dlx1 C-term.
1
105
dlx1 N-term.*
123
dlx2
1
270
dlx2
1
101
dlx2 N-term.
123
270
dlx2 C-term.
96
270
dlx2 C-term.*
137
dlx4
1
283
dlx4
1
107
dlx4 N-term.
77
283
dlx4 C-term.

Figure 3-2: Relative CAT activity of full length Dlx1 and its N- and C-terminal regions. Full length Dlx1 and Dlx1 fragments are represented schematically on the left. Numbers are as in figure 3-1. The relative CAT activity of the different constructs is shown on the right. The controls consist in 1) the expression vector PTL1 without any Dlx gene inserted; 2) the expression vector PTL1 with an inserted Gal-4 DBD. The greatest activation is observed with the fragment Dlx1 N-terminal* whereas the lowest activation is observed with the C-terminus fragment.

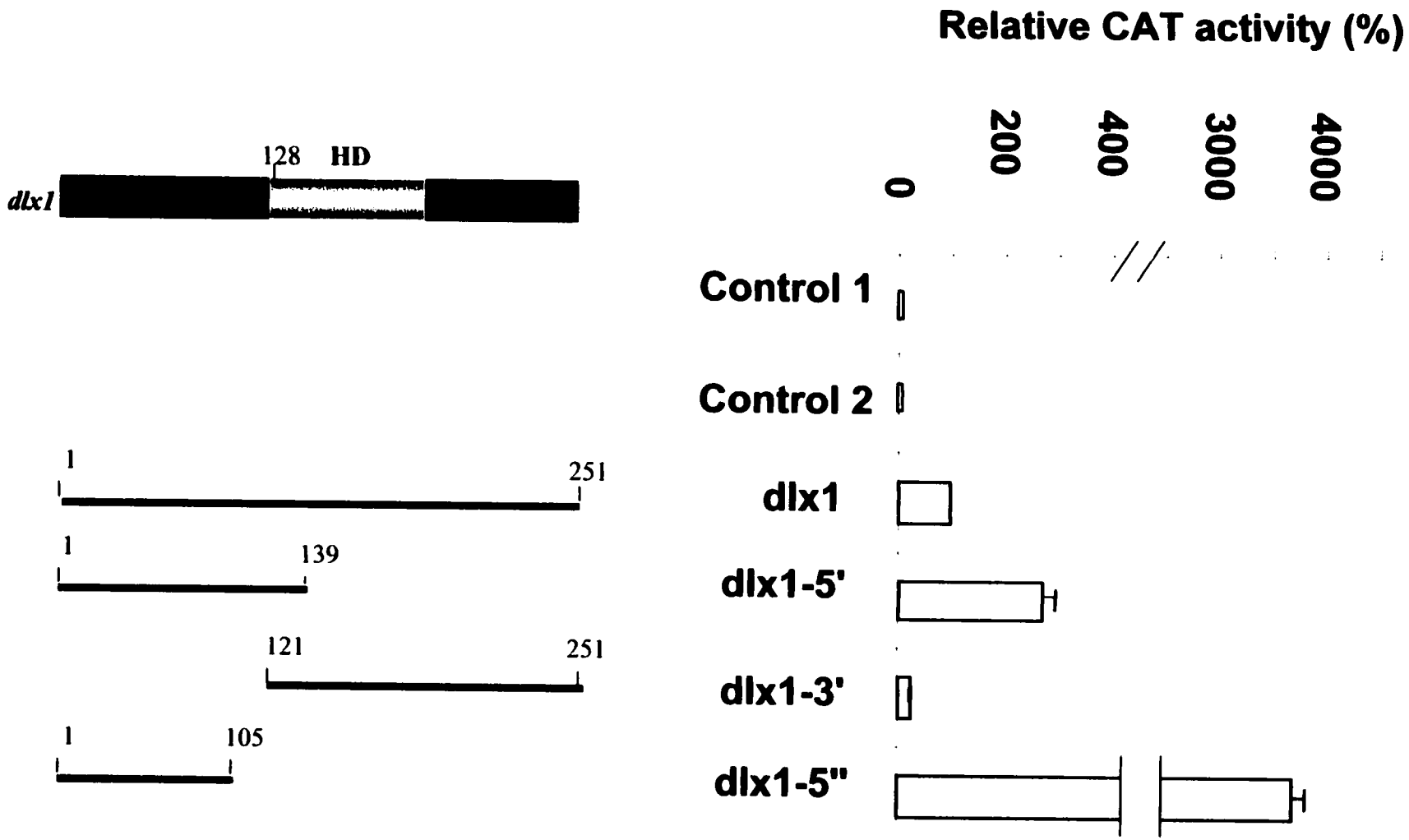

Relative CAT activity (%)
0
200
400
3000
4000
128
HD
dlx1
Control 1
Control 2
dlx1
dlx1-5'
dlx1-3'
dlx1-5"
1
251
1
139
121
251
1
105

Table 3-1:Relative CAT activity numbers of the different Dlx fragments

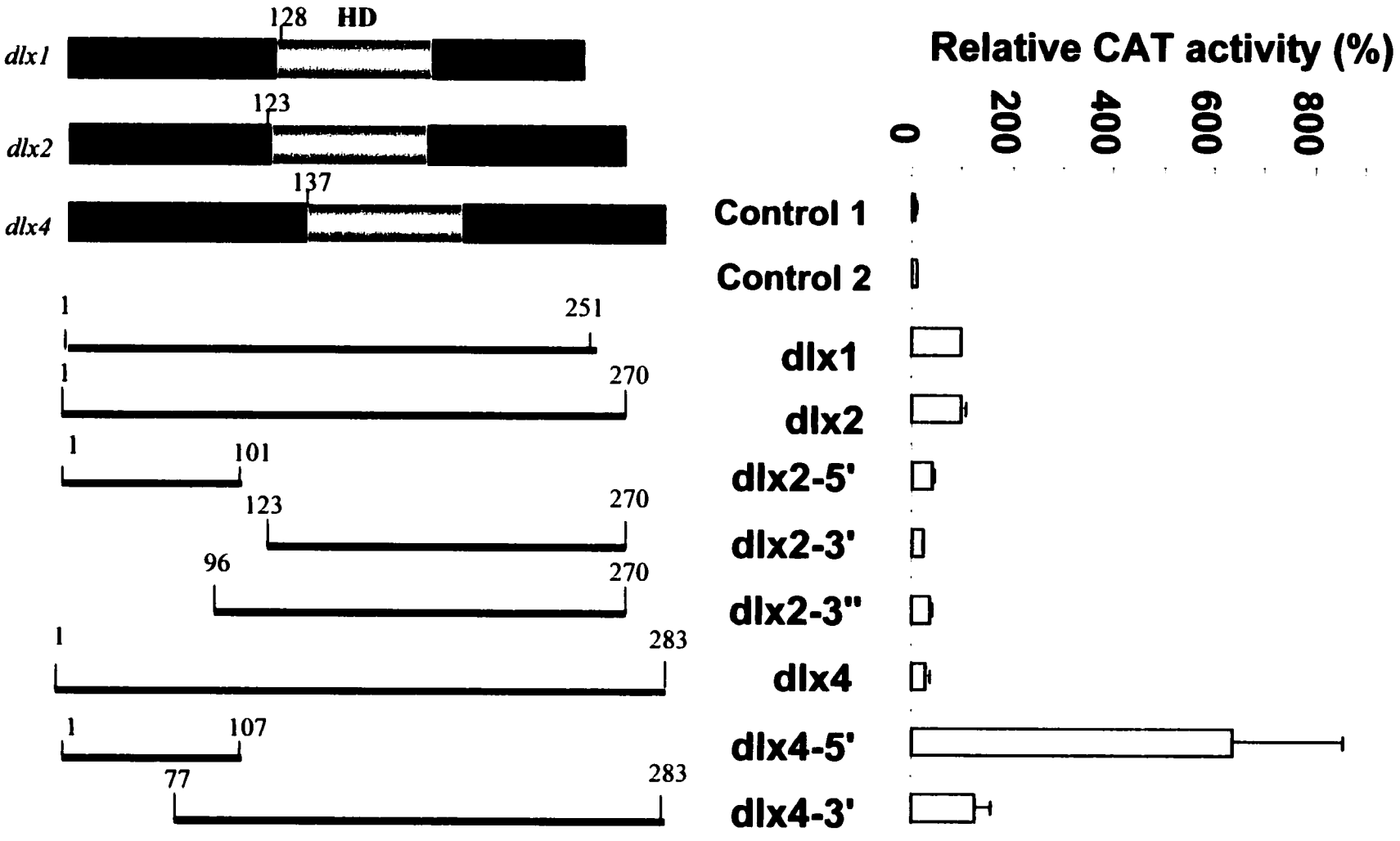
Values were calculated on the basis that full length Dlx1 activation is equal to 100%. The expression vector in which each Dlx fragment has been inserted is also mentioned as well as the length of the fragments, see also figure 3-1.

Fragments	%relative activation	Expression vector	amino acid sequence
Dlx1	100	PTL1	1 - 251
N-terminus	253 (±14)*	PTL1	1 - 139
C-terminus	18 (±3)	GalO	121-251
N-terminus*	3691 (±122)	PTL1	1 -105
Dlx2	98 (±9)	GalO	1 - 270
N-terminus	41 (±1)	PTL1	1 - 101
C-terminus	22 (±2)	PTL1	123 - 270
C-terminus*	36 (±5)	GalO	96 - 270
Dlx4	29 (±8)	PTL1	1 - 283
N-terminus	635 (±220)	GalO	1 - 107
C-terminus	125 (±34)	PTL1	77 - 283

* Standard error

Figure 3-3: Relative CAT activity of the full length Dlx2 and Dlx4 and their N- and C-terminal regions.

Full length Dlx2 and Dlx4 as well as their N- and C-terminal fragments are represented schematically on the left. Numbers are as in figure 3-1. The relative CAT activity of the different constructs is shown on the right. The controls consist in 1) the expression vector PTL1 without any *dlx* gene inserted; 2) the expression vector PTL1 with an inserted Gal-4 DBD. The greatest activation is seen with the Dlx4 N-terminus and the lowest with dlx2 C-terminal.


128
HD
dlx1
123
dlx2
137
dlx4
1
251
1
270
1
101
123
270
96
270
1
283
1
107
77
283
Relative CAT activity (%)
0
200
400
600
800
Control 1
Control 2
dlx1
dlx2
dlx2-5'
dlx2-3'
dlx2-3''
dlx4
dlx4-5'
dlx4-3'

This fragment activates transcription at high levels, 6 times that of full length Dlx1 and 2.5 times higher than the one observed with the N-terminus of Dlx1 (figure 3-3 and table 3-1). Surprisingly, the C-terminus of Dlx4 which contains 60 amino acids of the N-terminal region before the beginning of the homeodomain, is also a good activator (125% compared to Dlx1) (figure 3-3 and table 3-1). This contrasts with the other C-terminal fragments studied which produced considerably weaker reporter gene expression than the corresponding full length proteins (table 3-1).

Considering all these results we decided to look at the amino acid sequence alignment of the zebrafish Dlx fragments used in this study. We realised that the C-terminus of Dlx4 starts at amino acid 77 which is 60 aa before the homeodomain whereas Dlx1 and Dlx2 C-terminal fragments start at the beginning of the homeodomain. We also noticed that Dlx1 N-terminus spans a longer region than the N-terminal region of Dlx2 and Dlx4, meaning that the N-terminal of Dlx1 covers a 22 amino acid stretch located before the start of the homeodomain. This stretch is not found in the N-terminal regions of Dlx2 and Dlx4. Therefore, we decided to go further in the study and subclone additional fragments to get a better idea of where the activation domains are located. We shortened the N-terminus of Dlx1 and we also elongated the C-terminus end of Dlx2. The new Dlx1 N-terminus* is starting at aa 1 and ends at aa 105 which means this time that this fragment does not contains the 22 amino acid stretch located before the beginning of the homeodomain (figure 3-1). It has approximately the same length as the N-terminal fragments of Dlx4 and Dlx2 (figure 3-1). Surprisingly, this fragment activates transcription of the reporter even better. The activation level is 14.5 times higher than the previous N-terminus fragment of Dlx1 (figure 3-2 and table 3-1). Also, it is 6 times higher than the N-terminus of Dlx4 (figure 3-2 & 3-3 and table 3-1) even if they are about the same length. The new Dlx2 C-terminal* fragment starts this time at aa 96 which is at

27 aa in front of the homeodomain (figure 3-1). I previously mentioned that the Dlx2 C-terminus is a poor activator. The new fragment (C-terminus* of Dlx2) is showing an activation of 36% which is only 1.6 times higher than the previous C-terminal Dlx2 (figure 3-3).

3.1.2. Western blots

An essential control in these types of studies is to assess if all proteins or protein fragments were expressed at comparable levels. Nuclear extracts obtained from transfected cells were assayed by Western blotting with a monoclonal antibody against the GAl-4 DBD. This blot, presented in figure 3-4 is the result of two different experiments since for each transfection done for the western, we could not see all Dlx proteins. I was never obtaining a blot where all proteins or proteins fragments were detected. There were always two or three undetectable proteins. Therefore I had to use the remaining samples of two different experiments to do the final one. The antibody was able to detect almost all the proteins (figure 3-4 and table 3-2 for M.W.). The full length Dlx1 protein was hardly detected whereas the full length proteins of Dlx2 and Dlx4 were not. However, it does not mean automatically that these proteins are not expressed in CHO cells. It is probably a detection problem, since transcriptional activity is observed for these fragments. We observed also on every western blot, a band at lower molecular weight (~21.5Kd) (figure 3-4). The presence of this band could not be explained as well as the fainter upper bands detected on the blots.

Figure 3-4: Western blots of the Dlx-Gal-4 fusion proteins.

The Gal-4 DBD antibody was used to detect the fusion proteins. The intensity of the bands is proportional to the level of protein expression in transfected CHO cells. It appears that protein expression in the cells is equivalent except for Dlx1 N-terminal* and Dlx1 C-terminal for which the signal was stronger. Full length Dlx1 was barely detected whereas full length Dlx2 and Dlx4 were not detected at all.

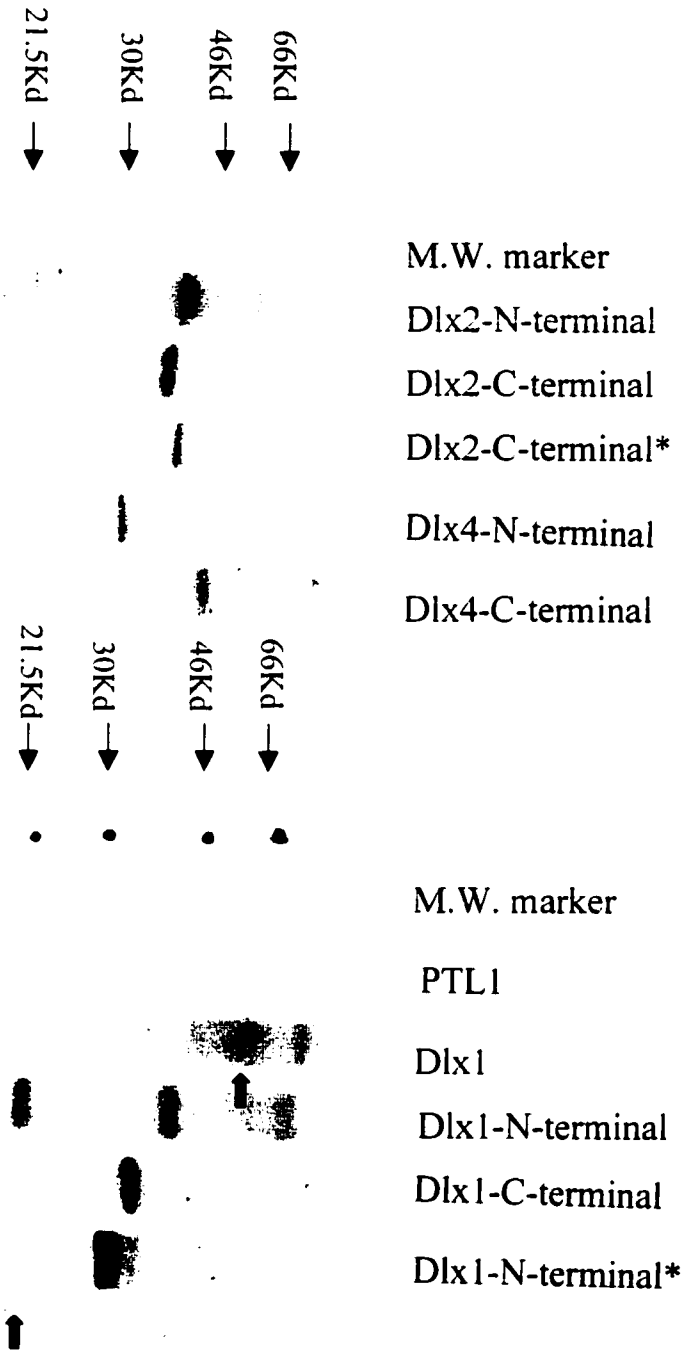

21.5Kd
30Kd
46Kd
66Kd
M.W. marker
Dlx2-N-terminal
Dlx2-C-terminal
Dlx2-C-terminal*
Dlx4-N-terminal
Dlx4-C-terminal
21.5Kd
30Kd
46Kd
66Kd
M.W. marker
PTL1
Dlx1
Dlx1-N-terminal
Dlx1-C-terminal
Dlx1-N-terminal*

Table 3-2: The predicted molecular weight of the different Dlx-Gal-4 DBD fusion proteins.

The molecular weight in Kda of full length Dlx1, Dlx2 and Dlx4 as well as their N- and C-terminal fragments fused to Gal-4 DBD was calculated on the basis that 1 amino acid = 110 da.

Dlx fragments+Gal-4 DBD	Protein M.W. (Kda) (1aa = 110dA)
Gal-4 DBD alone	16.3
Dlx1	47.0
Dlx1 N-term.	35.6
Dlx1 C-term.	33.8
Dlx1 N-term*	28.1
Dlx2	51.9
Dlx2 N-term.	35.9
Dlx2 C-term.	35
Dlx2 C-term.*	36
Dlx4	57.8
Dlx4 N-term.	31.7
Dlx4 C-term.	45.9

3.2. Loss or decrease of dlx4 expression after injections.

To understand the molecular mechanisms responsible for the overlapping patterns of *dlx* expression in zebrafish embryos, we examined the possibility that zebrafish *dlx* genes may cross-regulate each other. This hypothesis is based on the temporal expression patterns of the *dlx* genes, and the fact that this type of regulation is observed with other homeobox-containing gene families (Zappavigna et al. 1991; Arcioni et al. 1992; Faiella et al. 1994). It was shown previously by Zerucha et al (1997) that injection of truncated forms of *dlx3* mRNA and a chimera consisting of the amino terminal half of *dlx3* and the homeodomain and the carboxy terminal half of *dlx2* mRNA, perturbed endogenous *dlx4* expression (Zerucha et al., 1997). However the percentages of the embryos for which they observed a reduction or a loss of *dlx4* expression were low (from 16% to 38%). It was imperative to improve the conditions of this experiment. We decided to insert the cDNA's of full length *dlx3*, truncated *dlx3* and the chimera *dlx3-dlx2* in the CS2+ vector instead of bluescript, which was used in the experiment described above. In contrast to bluescript, the CS2+ vector contains an SV40 poly A site which confers more stability to the synthesized mRNA (figure 2-5). Thus it is expected to see more severe phenotypes in a greater percentage of the embryos than what it was observed in the previous experiment. Similarly to the previous experiment, a full length and a truncated version of *dlx3* were ubiquitously expressed in zebrafish embryos by injecting in vitro transcribed mRNA coding for the transcript into single-cell embryos, and the expression of *dlx4* was examined by *in situ* hybridization 27 hrs later (data not shown) (figure 3-5). Further, a chimera construct consisting of the amino-terminal half of Dlx3 (from position 1 until the 9th amino acid of the homeodomain) and the homeodomain and carboxy-terminal half of Dlx2 (from the same position to the carboxy terminus) was injected in order to determine

if the carboxy-terminal half of Dlx2 is able to functionally substitute for the same domain of Dlx3 (figure 3-6).

Expression of *dlx4* was only perturbed significantly in embryos in which the Dlx3-Dlx2 chimera is ubiquitously expressed. Of 330 embryos injected with a 0.5 ug/ul solution, approximately 63% of them died. In the 122 embryos that survived, a variety of phenotypes were observed. Malformations of the whole body were the most obvious phenotype observed. In the remaining live embryos, approximately 30% of them were severely affected morphologically (shorter median finfold). The phenotypes related to *dlx4* expression that were the most frequently encountered are: a loss or a reduction of *dlx4* expression in the ear, in the forebrain and in the visceral arches in 13% of them (figure 3-7 C,D and figure 3-8 B and table 3-3). A loss or a reduction was observed only in the visceral arches in 12% of injected embryos (figure 3-7 E-F and table 3-3). A loss or a reduction of *dlx4* expression in the ear and the visceral arches only in 11% of injected embryos was also observed (figure 3-8 C and table 3-3). Finally, in 9% of injected embryos showed a loss or a reduction of *dlx4* expression in all the regions where *dlx4* is normally expressed (ears, visceral arches, forebrain and median finfold). It is surprising that no expression is observed in the median finfold, since the results obtained by Zerucha et al. showed that *dlx4* expression was never reduced or lost in that region of the embryo. Additional phenotypes related to *dlx4* expression were observed albeit not frequently (table 3-3). The total percentage of embryos affected by the injection of the chimera of dlx3-2 is 83% including those who died. If the dead embryos are excluded (29%), the percentage of living embryos affected is 54% which is a significant result considering that all the embryos injected with a control which consisted of injecting a mixture without RNA were normal and the percentage of those who died was considerably lower (~12%).

Figure 3-5: Schematic representation of the Dlx injection strategy.

This figure summarizes the experiments done in zebrafish embryos. Injections of mRNA were done at the one cell stage. The injected embryos were left to develop for 27 hours and then fixed and hybridized to a *dlx4* probe to determined if those injections perturbed Dlx4 expression. The picture of a 24h zebrafish embryo sagittal view, dorsal to the top, anterior to the left, at the bottom left, illustrate the wild-type expression of *dlx4* which is restricted to the forebrain (F), otic vesicle (OV), visceral arches (VA) and the median finfold (MF).

On the right, the injected embryos were left to develop for 5 days and then fixed with 4% PFA. To see if ubiquitous expression of *dlx* mRNA affected the formation of the craniofacial cartilage elements, which are visible at this stage, alcian blue staining was used. A picture of a 5 days old larvae ventral view, anterior to the left shows the stained craniofacial cartilage.

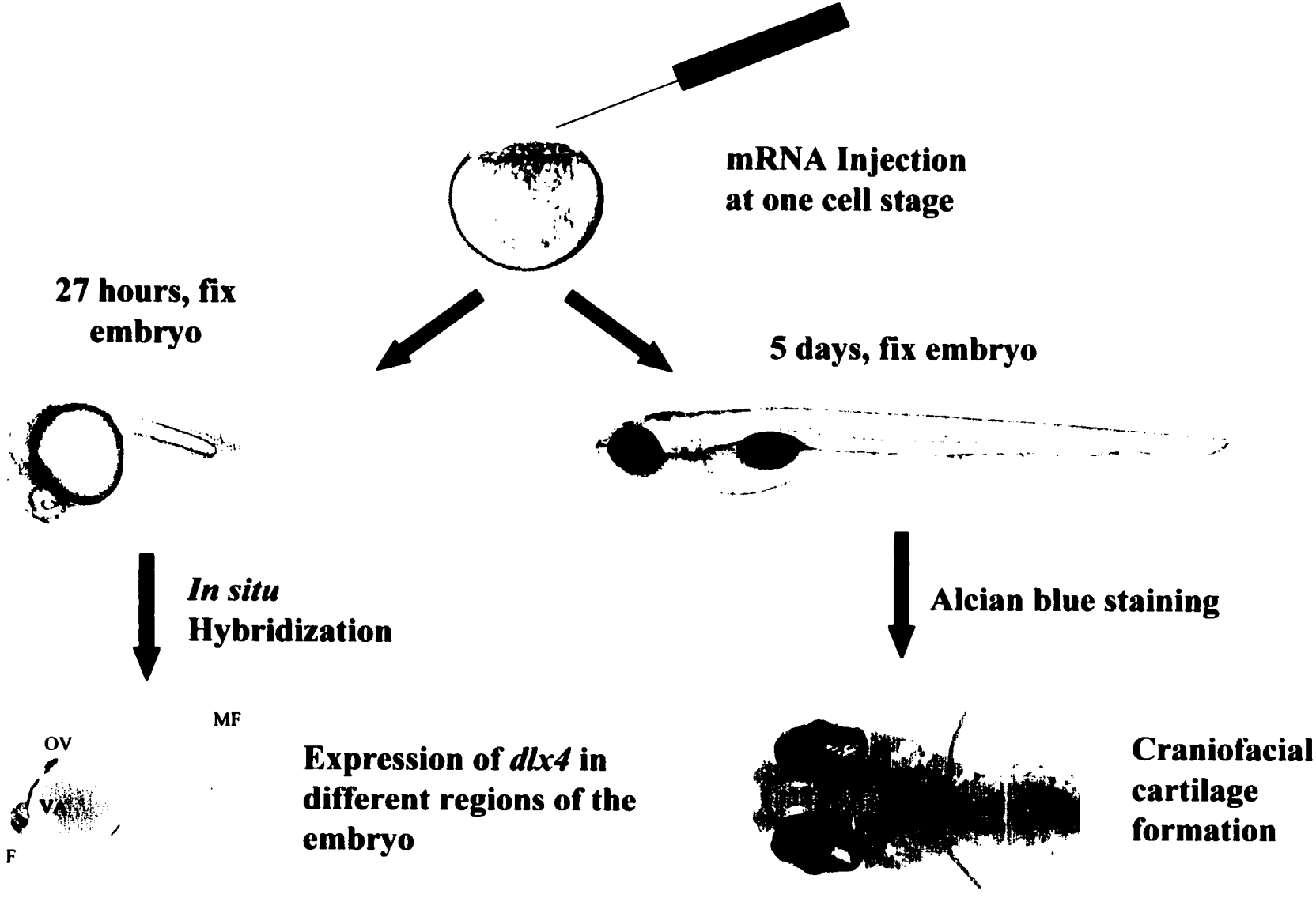

mRNA Injection
at one cell stage
27 hours, fix
embryo
5 days, fix embryo
In situ
Hybridization
Alcian blue staining
MF
OV
VA
F
Expression of dlx4 in
different regions of the
embryo
Craniofacial
cartilage
formation

Figure 3-6: mRNA constructs.

Summary of synthetic mRNA coding for full length, truncated and chimeric *dlx3* and full length and truncated *dlx2* used for injections in zebrafish embryos. The restriction site used to create the truncated *dlx3* and *dlx2* is XhoI. The truncated *dlx3* mRNA codes for a polypeptide of 131 amino acids in length consisting of the amino terminal region of Dlx3 as well as the first nine amino acids of the homeodomain. The truncated *dlx2* mRNA codes for a polypeptide of 131 amino acids in length consisting of the amino terminal region of Dlx2 as well as the first nine amino acids of the homeodomain. The chimera was created using the conserved XhoI site in the homeoboxes of *dlx2* and *dlx3*. The creation of this chimera does not affect the reading frame of *dlx2* nor *dlx3*. The chimeric mRNA codes for a polypeptide consisting of the amino terminal fragment and the first nine amino acids of the homeodomain of Dlx3 and the carboxy terminal fragment and most of the Dlx2 homeodomain.

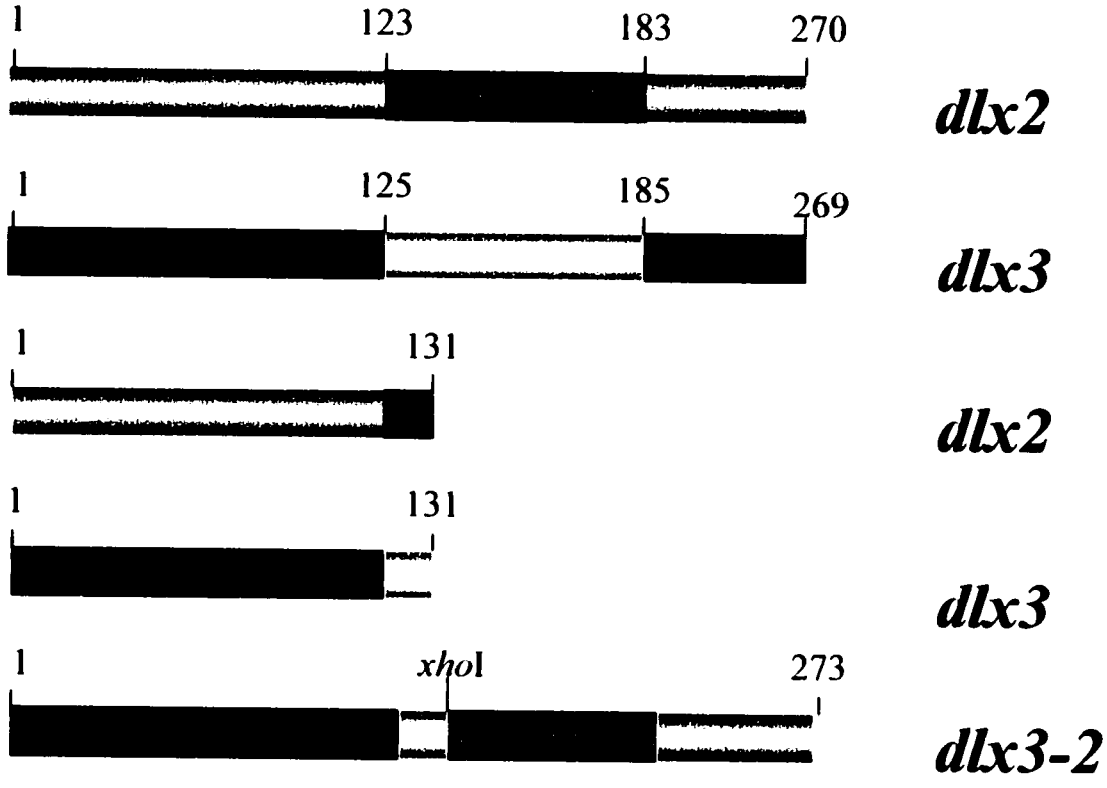

1
123
183
270
dlx2
1
125
185
269
dlx3
1
131
dlx2
1
131
dlx3
1
xhoI
273
dlx3-2

Figure 3-7: Injection of the Dlx3-Dlx2 chimera perturbed *dlx4* expression in zebrafish embryos.

Perturbations in expression of *dlx4* as a result of the injection of the Dlx3-Dlx2 chimera: sagittal (A, C and E) and dorsal (B, D and F) views of *dlx4* expression in the forebrain (F), visceral arches (VA), and otic vesicle (OV) of control embryos. The control embryos were injected with a saline solution. (A-B) wild-type *dlx4* expression. (C-D) A loss or reduction of *dlx4* expression was observed in the otic vesicle, in the forebrain and in the visceral arches in 13% of the embryos. (E-F) A loss or reduction of *dlx4* expression was observed only in the visceral arches in 12% of the injected embryos.

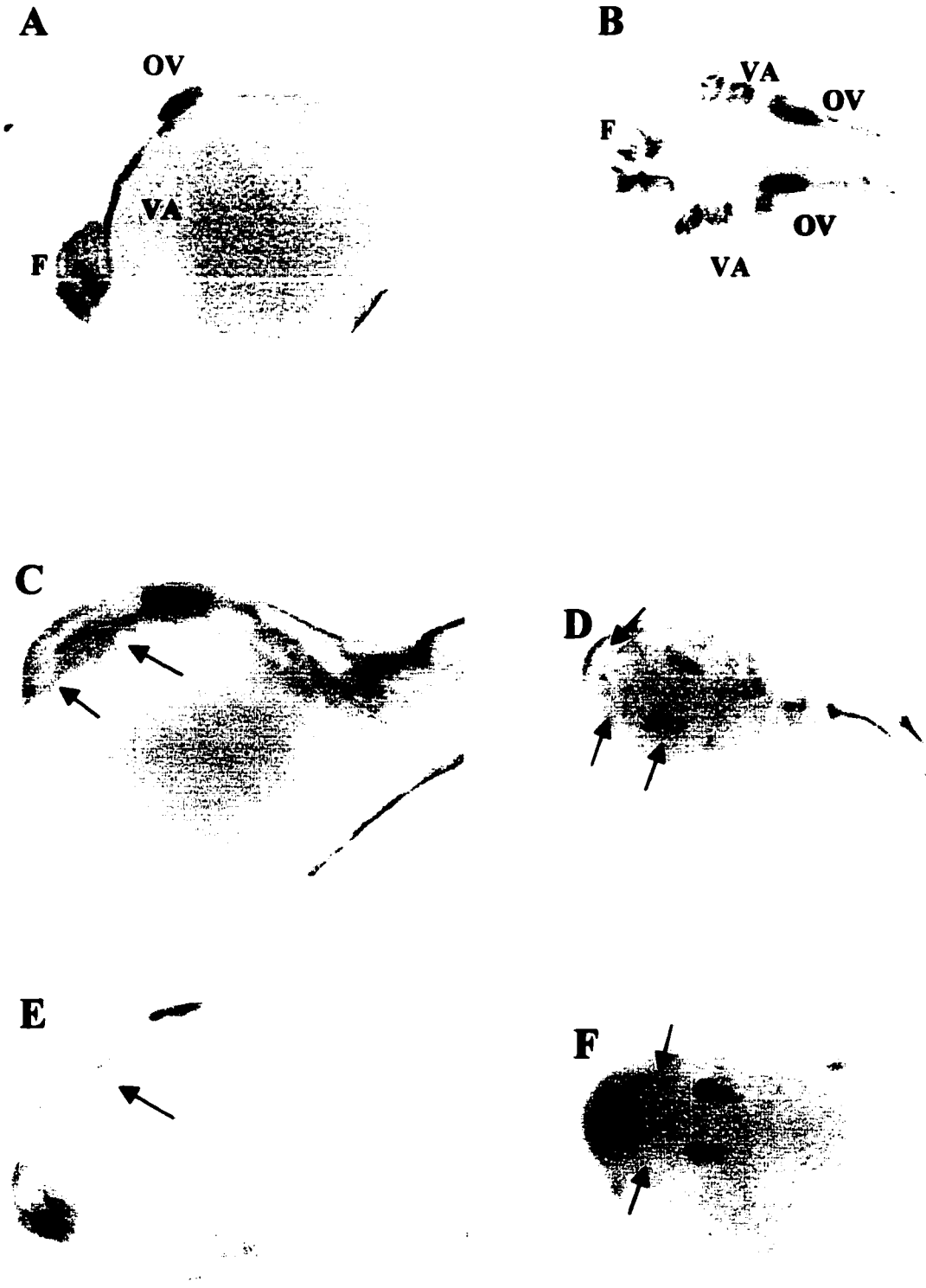

A
OV
VA
F
B
VA
OV
F
OV
VA
C
D
E
F

Figure 3-8:Perturbations of *dlx4* expression as a result of the chimera Dlx3-Dlx2 injection.

Perturbations in *dlx4* expression as a result of the injection of the Dlx3-Dlx2 chimera. (A) Wild-type *dlx4* expression. (B) Another embryo picture of the 13% ones that showed a loss or a reduction of *dlx4* expression in the otic vesicle, in the forebrain and in the visceral arches. (C) A loss or reduction of *dlx4* expression was observed in the otic vesicle and the visceral arches in 11% of the embryos injected.

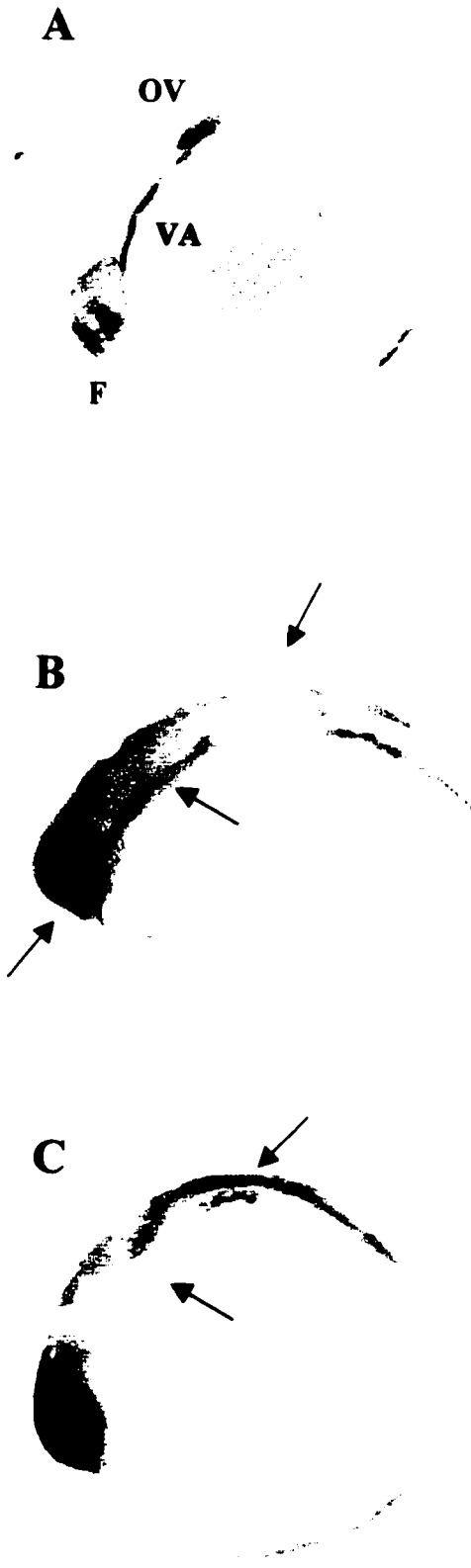

A
OV
VA
F
B
C

Table 3-3: Phenotypes related to *dlx4* expression observed after injections of the *dlx3-dlx2* chimera.

This table lists the various phenotypes related to *dlx4* expression obtained after injection of the Dlx3-Dlx2 chimera. *In situ* hybridization was used to detect endogenous *dlx4* mRNA in 27 hour embryos. The most frequently observed phenotypes are: a loss or reduction (indicated by a -) of endogenous *dlx4* in the visceral arches in 12% of the injected embryos, also a loss or reduction of *dlx4* expression in the ears, forebrain and visceral arches in 13% of them and a loss or reduction of *dlx4* expression in the ears and the visceral arches in 11% of the embryos.

mRNA dlx2-3	Ears	Forebrain	Visceral Arches	Finfold	% embryos
	+	+	+	+	49
	+	+	-	+	12
	-	-	-	+	13
	-	+	-	+	11
	-	-	-	-	9
	+	-	+	+	2
	+	-	-	+	2
	-	+	-	-	2
	-	+	-	-	1

Injections of the full length *dlx3* and *dlx2* were done to verify if their overexpression would also perturb *dlx4* expression. A small percentage of embryos were affected by the injection of these two constructs (17% and 16% respectively). The phenotypes observed were severe malformations of the whole embryo in ~30% of the embryos (shorter median finfold and smaller eyes), showing a loss or reduction of *dlx4* in some region (the ear, the visceral arches and the forebrain) of the embryos. The injections of the complete *dlx2* and *dlx3* that were done previously by Zerucha et al. did not affect the *dlx4* expression or the morphology of the embryos. Injections of relatively high concentration of mRNA and increased mRNA stability confered by CS2+ vector is probably the cause of the phenotypes that we observed.

Other injections of synthetic mRNA coding for truncated forms of *dlx3* and *dlx2* proteins that end at the carboxy terminal portion of the homeodomain were done (figure 3-6). Only 5% of the embryos showed a loss or reduction of the *dlx4* expression in the ear, the visceral arches and the forebrain, when the truncated form of *dlx3* was injected. This percentage of affected embryos is lower than the one observed when full length *dlx3* is overexpressed. The percentage of the dead embryos after 27 h of development was comparable to what it is observed with the control embryos. A larger percentage of embryos were affected when the truncated *dlx2* was injected. Effectively 19% of the embryos injected presented a loss or a reduction of *dlx4* expression in the visceral arches, in the ears and in the forebrain. Also, approximately 30% of the embryos were morphologically affected by the overexpression of truncated *dlx2*.

3.3. Craniofacial dysmorphogenesis after injection.

The injection of the chimeric version of *dlx3-2* perturbed the *dlx4* expression in the visceral arches, the site of craniofacial cartilage morphogenesis. To see if the loss or reduce *dlx4* expression in the visceral arches affected the morphogenesis of the craniofacial cartilage structures which are visible 5 days post-fertilisation, injection of the chimera Dlx3-Dlx2 was done at one cell stage. To be able to detect any dysmorphologies in the cartilage, alcian blue staining was used to visualize it (figure3-5 right panel).

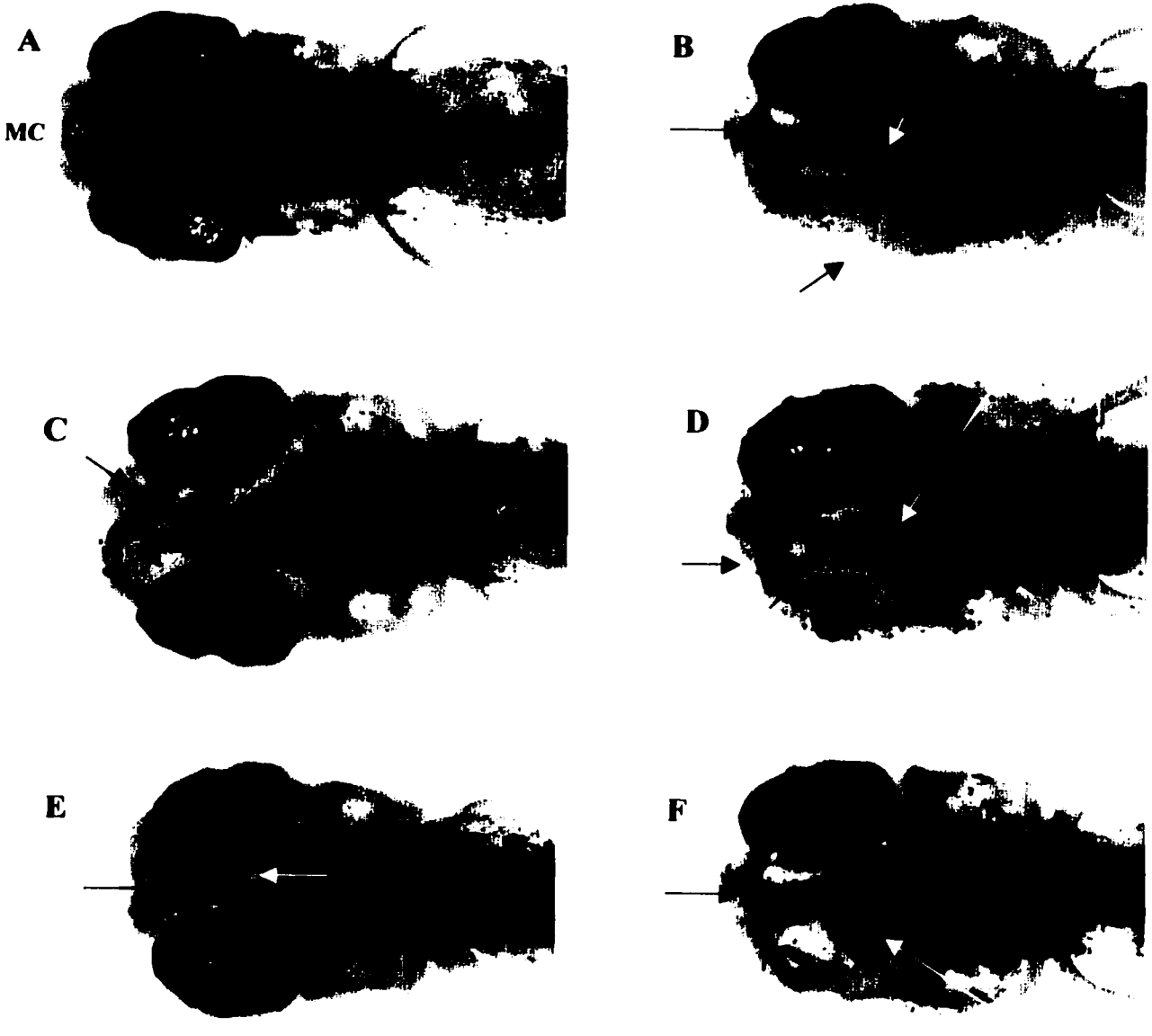
The same injection procedure as described above was used here. The embryos were injected with a 0.5 μg/μl solution of the chimera and with all other constructs. Only the Dlx3-Dlx2 chimera produced an effect on the cartilage. Again more than half of the embryos died before 24 h. Of the 121 embryos that survived, 35% of them were seriously malformed (figure 3-9 B-C). The phenotype that we see more often is an eye missing in 13% of the embryos (figure 10 B-D-F). The Meckel's cartilage, the Trabeculae, the Basihyal and the Ceratohyal cartilage were malformed in 6% of the embryos (figure 3-10 B-F). The Meckels cartilage shape was larger compared to control (figure 3-10 B-D-F) and in some of them its normal U shape was not well formed (figure 3-10 C-E). The ceratohyal cartilage did not have the characteristical V shape, as shown in control larvae anymore (figure 3-10 C-E-F). Also in some of the larvae, its morphogenesis appeared incomplete (figure 3-10 B-D). In some of the fishes, the basihyal was not detectable (figure 3-10 B-C) whereas in other ones, it was missplaced i.e. not at the junction of the ceratohyal (figure 3-10 D). The trabeculae which normally have a V shape, seemed to be fused together at their anterior extremities, as shown in figure 3-10 B-F). In addition, a large percentage of embryos (11%) showed a curved tail (figure 3-9 D).

Figure 3-9: Malformations of 5 day old larvae after injection of the Dlx3-Dlx2 chimera. This figure shows the typical malformations observed in 5 days old larvae as a result of the Dlx3-Dlx2 overexpression. (A) Sagittal view of control larvae injected with a saline solution. (B-C) Two examples of larvae severely affected morphologically; 35% of the embryos injected showed smaller eyes and shorter tails compared to controls. (D) In 11% of the injected embryos, a side-curved tail phenotype was observed.


A
B
C
D

Figure 3-10: Ubiquitous expression of the Dlx3-Dlx2 chimera caused dysmorphogenesis of the craniofacial cartilage.

Dysmorphologies of 5 day old larvae craniofacial cartilage was observed when the Dlx3-Dlx2 chimera was injected. Ventral view (anterior to the left) of craniofacial cartilage: Meckels cartilage (MC), trabeculae (T), basihyal (B) and ceratohyal (C). (A) Control larvae craniofacial cartilage stained with alcian blue. (B-D and F) The Meckels cartilage shape was larger compared to control. (C and E) In some of them the normal U shape was not well formed. (C-E and F) The ceratohyal cartilage (arrow) did not present the characteristical V shape, as shown in control larvae. (B-D) In some larvae the ceratohyal morphogenesis seemed not to be complete. (B-C) In some of the larvae, the basihyal was not detectable. (D) In other ones the basihyal was missplaced i.e. not formed next to the junction of the ceratohyal. (B-F) The trabeculae which normally have a V shape, seemed to be fused together at their anterior extremities.


A
MC
B
C
D
E
F

The severity of the phenotypes is proportional to the concentration of RNA injected. A smaller concentration (0.1 μg/μl) was injected to see if the same malformations would be observed after 5 days of development. All embryos were normal and the proportion of dead embryos at 27 h dropped to 17% which is only slightly higher than that seen in control embryos (12%) (table 3-4). The injection of a 0.5 μg/μl solution caused the death of 63% of the embryos after 24-27 h and 29% of the remaining live embryos were morphologically affected (table 3-4). The injection of a 1μg/μl solution had devastating effects, with 93% dead embryos (table 3-4). The volume of injection was approximately always the same throughout the experiment (10 nl) and those test injections were done on the same day.

3.4. Morphology of the inner ear after injection of the chimera

Since *dlx4* expression is disturbed in the inner ear for 24% of the embryos injected with the chimera, the morphology of this structure was verified with a confocal microscope. The embryos were injected with the *dlx3-dlx2* chimera and the embryos fixed at 5 days. The actin filaments of the embryos were then colored with alexia 488 phalloidin to detect the hair cells of the inner ear and look for any abnormalities in their formation. Also, the other structures such as the otic capsule, the vestibular ducts, the maculae and the cristae were examined.

The structure of the inner ear of injected fish was examined at 5 days. All the inner ear structures seemed to have developed normally. We were able to see by confocal microscopy the presence of the hair cells aligned on the cristae (figure 3-11 C, F) and the maculae (figure 3-11 B, D, E). The cristae and the maculae were well formed and they were all present; three cristae (figure 3-11 C, F) and two maculae (figure 3-11 B, D, E).

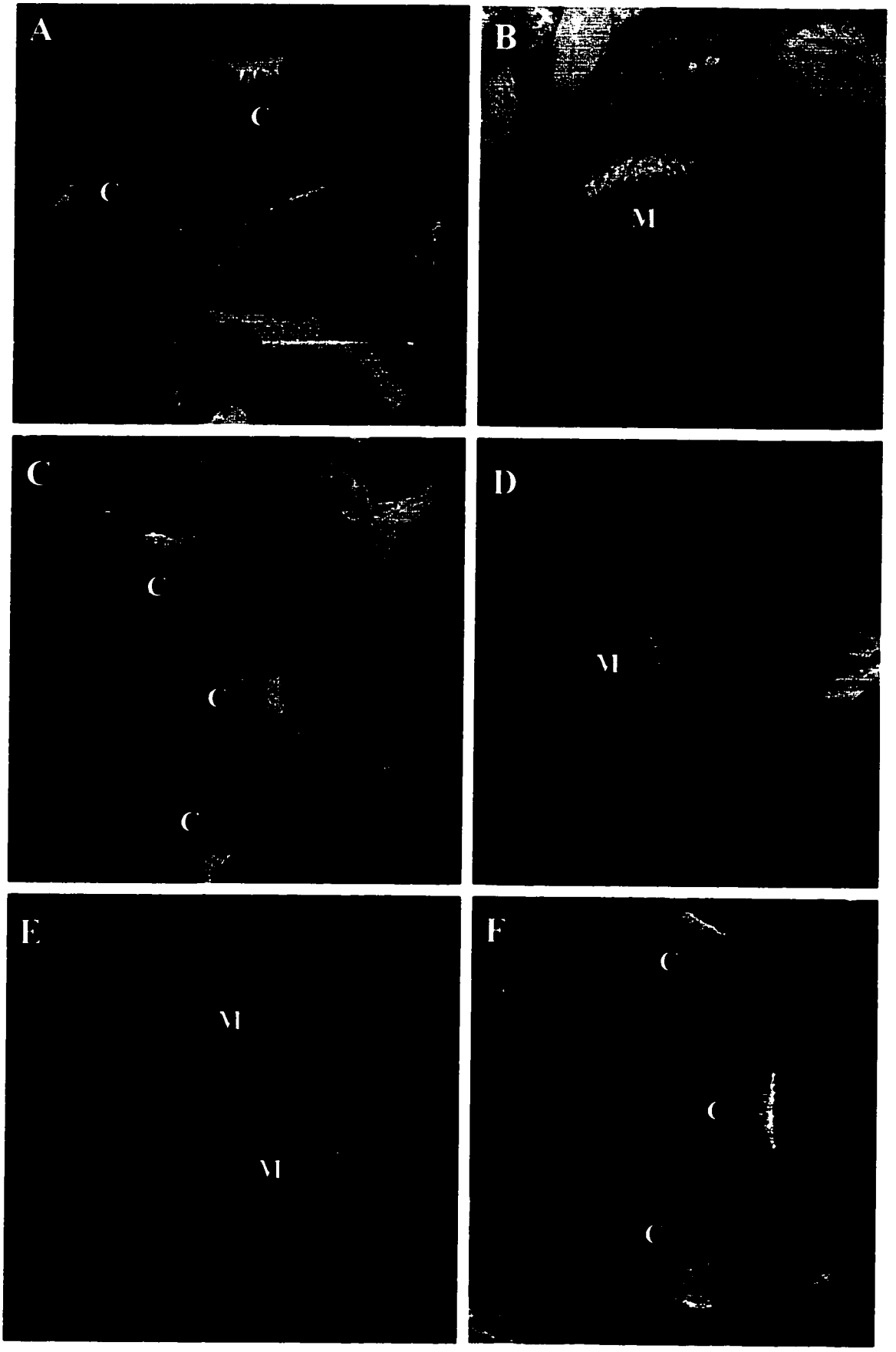
Table 3-4: Effects of the concentration of injected mRNA on embryonic phenotype.

To verify if the concentration of mRNA injected was influencing the normal development of embryos, a series of injections of increasing concentration of Dlx3-Dlx2 chimera were done. Control embryos were injected with a saline solution. Embryos were injected with increasing concentrations of the chimera mRNA. Only 12% of the embryos injected with saline solution died, whereas injection of increased concentration of mRNA increased also the percentage of dead embryos. Injection of 0.1μg/μl mRNA solution produced 17% of dead embryos, which is still a low percentage when compared to controls, whereas the injection of 0.5μg/μl of mRNA solution produced 63% of dead embryos and 29% of malformed embryos (smaller eyes and shorter tail). Injection of 1μg/μl of mRNA solution almost killed the totality of the embryos (92%).

	Control (no mRNA)	0.1 ug/ul	0.5 ug/ul	1 ug/ul
# embryos total	26	42	49	34
# malformed embryos	0	0	14 (29%)	1 (8%)
# dead embryos	3 (12%)	7 (17%)	31 (63%)	33 (92%)

Figure 3-11: Inner ear morphology after injection of the Dlx3-Dlx2 chimera.

The inner ear structures like the cristae (C) and the maculae (M) were not affected by the ubiquitous expression of the chimera. (A) Confocal microscope produced the image illustrating two cristae well formed in the inner ear of a control larvae as well as the presence of hair cell bundles at the tip of the cristae, stained with alexia 488 (green color). (B-F) Structures of the inner ear of injected embryos. (B-D) We see a well formed maculae with the hair cell bundles stained. (C and F) The three cristae of the inner ear of injected fishes are shown. Their structure seemed normal as well as the formation of hair cells as shown by the colored ciliae. (E) A view of two maculae that seemed well formed and the presence of hair cell ciliae stained by the alexia 488.


A
C
C
B
M
C
C
C
C
D
M
E
M
M
F
C
C
C

We saw also the vestibular ducts (figure 3-11 C). Thus the morphology of the hair cells and the other region of the inner ear were totally normal for all embryos examined. Even the most severely affected fishes like the one on figure 10 B, C showed normal inner ear structures. Thus, injections of the chimera did not seem to perturb inner ear development.

4. Discussion

4.1. Probable location of the activation domain.

Transfection experiments were done in order to identify domains responsible for transcriptional activation. The *dlx* gene fragments studied (*dlx1*, *dlx2* and *dlx4*) had different levels of activation. The strongest activation is seen with the fragments that represent the most 5' region of the genes *dlx1* (Dlx1 N-terminus*) and *dlx4*. The lengths of the amino acid sequence of these two fragments are comparable, that is, 105 amino acid long for Dlx1 and 107 for Dlx4. The amino acid sequences of these regions are different suggesting that the activation domain sequence is not a consensus sequence.

One possibility is that the activation domain depends only on a few amino acids. Those amino acids could confer a tri-dimensional structure to the protein which is favourable for the activation of the transcription. Shirasawa et al., analysed the primary structure of rat Dll-3 gene and compared it with published Dlx homeoproteins of different species, including the zebrafish Dlx3 which share 46.1% identity with zebrafish Dlx3 (Ekker et al. 1992). They found for the first time a consensus motif of 18 amino acid residues which was located in the N-terminal region and well conserved among Dll members. This sequence appears to be highly specific for Dll homeodomain proteins, since it was not found in homeoproteins of other classes nor in other proteins (Shirasawa et al., 1994). They also found another consensus motif in which 8-11 tyrosine residues appear within 50 amino acids. Some of these tyrosine residues are separated by an interval of 7-8 amino acid residues or 14-15 amino acid residues, which would fit in 2 or 4 turns of a helix structure analogous to the leucine zipper. This motif could allow protein-protein interactions. Shirasawa and al suggest that the two consensus motifs in the N-terminal region of Distal-less family may be involved in protein-protein interaction

necessary for transcriptional activation. So it could be that one of these two domains corresponds to the activation domain that we found in the N-terminal region of Dlx1 and Dlx4. It is also reinforced by the results obtained with the 3' region or C-terminus of Dlx1 and Dlx2, which do not activate transcription at levels as high as the N-terminal region. It suggests that transcriptional activation is mainly conferred by the N-terminal region.

The N-terminal region which spans the whole amino acid sequence before the start of the homeodomain seems to contain, in addition to an activation domain, either a repressor domain or some sequences with negative effects on transcriptional activation at least in the context of a Gal-4 DBD fusion protein. Effectively, the activity of Dlx1 N-terminus which spans the whole region and stops at the 12^{th} amino acid of the homeodomain (aa position 139) is lower than the one observed for the shorter Dlx1 N-terminal* that ends at 22 aa before the homeodomain. Similar results were obtained with Dlx4 fragments. The activity of the N-terminal fragment which stops at 28 aa before the homeodomain, is very high whereas the C-terminal fragment which contains 60 aa before the homeodomain still showed high activation compared to full length but lower than the N-terminus fragment by 5 fold. Initially, it looks like this fragment contains an activation domain since it activates transcription almost as high as the longer Dlx1 N-terminus fragment (125% versus 253%). Since it was shown that Dlx1 and Dlx2 C-terminal fragments which start just at the beginning of the homeodomain and span the whole C-terminal region, activate transcription at lower levels (~6-7 fold lower than Dlx4 C-terminus), this suggested that maybe the stretch of 60 amino acids before the homeodomain present in Dlx4 C-terminal fragment and absent in Dlx2 and Dlx1 C-terminus, is able to activate transcription. However, the removal of a stretch of 22 aa before the homeodomain of Dlx1 N-terminal region resulted in a much higher activation (14.5 fold higher than Dlx1 N-terminus activity). This suggested to us finally, that this small region among *dlx* genes before the homeodomain may contain sequences that

would have a negative effect on transcriptional activation. That would explain in part, why the N-terminal region of Dlx1 and Dlx4 that do not contain the region before the homeodomain showed higher activity than the complete Dlx1 and Dlx4. In summary the results suggested that the activation domain is located in the N-terminus and that a domain containing sequences that would have a negative effect on transcriptional activation would be at the junction between the N-terminus and the homeodomain. In addition to this domain with negative effects on transcriptional activation, another explanation is possible to explain why the full length Dlx proteins activated less the transcription than the N-terminus alone of Dlx1 and Dlx4. It is maybe because of the presence of 2 motifs found by Shirasawa et al in the C-terminus region of Dll homeoproteins that are serine rich and that can be phosphorylated. Shirasawa et al suggested that the serine rich motifs in the C-terminal region may be involved in some modifications of the Dll homeoprotein. Furthermore, Dong et al suggested that phosphorylation of homeoproteins may be a general mechanism for modulating homeodomain protein activity. They concluded that phosphorylation of a threonine 263 in the *Drosophila* Ftz protein is crucial for its function in developing embryos and it is likely to be phosphorylated when the protein is active (Dong et al., 1998). Similarly to the homeodomain Ftz protein, phosphorylation of the serine-rich motifs found in Dll homeoproteins could influence the activity of the whole protein during development. It is possible that those motifs once phosphorylated are involved in silencing protein activation. The removal of the Dlx C-terminus would no longer be able to influence the activity of the protein. This could explain the higher activation of the N-terminus fragment observed for Dlx1 and Dlx4 versus the complete proteins.

The stability of the full length proteins could also explain the lower activation. It can be possible that complete Dlx1 and Dlx4 are less stable in the cells than the N-terminal fragments as seen on the western blots. The Gal-4 DBD antibody was not able to

detect the Gal-4-full length Dlx4 and Dlx2 fusion proteins and was barely detecting Gal-4-full length Dlx1. The possible instability of Dlx1 and Dlx4 full length proteins could be attributed to the presence of the fused Gal-4 DBD. The stability of the Gal-4-N- and C-terminal fragments would not be affected by the presence of the Gal-4 DBD. The Gal-4 DBD would confer a tri-dimensional conformation of the full length Dlx proteins not favourable for activating transcription compared to the fused Gal-4 DBD-Dlx fragments. It could be also simply be a detection problem by the Gal-4 DBD antibody. Dlx2 full length, in contrary was shown to activate more than its corresponding N-terminal fragment. However Dlx2 protein was not detected by Gal-4 DBD antibody. It may be due to a problem related to the western technique or Dlx2 might function differently than Dlx1 and Dlx4 even though it contains the two characteristic motifs in the N-terminal and C-terminal region found in Dlx members.

The tri-dimensional configuration of the Dlx proteins fused to the Gal-4 DBD may influence the level of transcriptional activation of the different Dlx fragments. It would be interesting to test the level of activation of the different Dlx fragments without the Gal-4 DBD using a natural Dlx DNA-response element. This is now possible since a student in the lab. T. Zerucha, isolated conserved sequences in the intergenic region that separates pairs of *dlx* genes (Ellies et al 1997a), which contains the ATTA/TAAT core. This student also tested the ability of a few Dlx proteins (the full length) to activate transcription from these elements in transient transfection assays.

Vigano et al studied the transcriptional activity of the Hox genes (Hoxd9, Hoxb1 and Hoxb3) using two different experimental strategies in order to locate their activation domains. They observed different results depending of the method used. For example, they studied the activation domains by co-transfecting the full length N-terminus and deleted N-terminus fragments with a luciferase construct containing a Hox control region.

They observed that a deletion of 75 aa at the beginning of the N-terminus, did not affect the transcriptional activity. However, larger deletion greatly reduced the transcriptional activity. They decided to use another method, similar to the one we used in our studies, which consists of creating a chimera between these Hox fragments and the Gal-4 DBD. The Gal-4 DBD is also, in this case located upstream the Hox fragments. They co-transfected these constructs with a luciferase reporter containing a 5-mer Gal-4 responsive element. They showed that the transcriptional activity of the same deletion fragment (75 aa) is reduced by 80% compared to the non-deleted N-terminus, whereas in the previous method the activity is as high as the full length protein. The transcriptional activity of the fragment with larger deletions (142 aa, 222 aa and complete N-terminus), showed the same activity with both methods. Therefore, it is reasonable to think that the Dlx fragments used in this study could show differences in transcriptional activity, depending on whether they are used in their normal context or when used as fusion proteins with the Gal-4 DBD.

It would also be interesting to delete the extremities of the dlx 5' and 3' regions of the Dlx proteins. Feledy et al. studied the transcriptional activation of the protein Distal-less 3 in *Xenopus* by progressively deleting the 5' and 3' extremities of the gene (Feledy et al., 1999). They found that expression of Dlx3 protein in either HeLa cells or *Xenopus* embryos resulted in strong activation of a model target gene construct containing three tandem copies of a Dlx3 binding site upstream from the TATA element. In addition, they found, by deletion analysis that Dlx3 has two activation domains; one located in the N-terminus region and the other in the C-terminus region.

Interactions with other proteins are also a factor that might influence transcriptional activation. Zhang, H. et al, demonstrated that dimerization of Msx and Dlx which occur through their homeodomain, displayed reciprocal inhibition. Specifically,

Msx proteins act as transcriptional repressors and Dlx proteins act as transcriptional activators, while in combination, Msx and Dlx proteins counteract each other's transcriptional activities (Zhang et al. 1997). It would also be a good idea to test the Dlx activity in presence of other proteins which have been shown to interact with them, such as the Msx proteins.

The Dlx protein fragments showed different levels of transcriptional activation. Thus it is possible, that the Dlx proteins play a specific role in regulating transcriptional activity. We could imagine a Dlx protein that would be a strong activator allowing it to regulate specific targets during development. Other Dlx proteins would be weaker activators and exhibit different functions than the stronger one. To elucidate the role in transcriptional activity of Dlx proteins, different methods, like the one described above have to be done.

4.2. The dlx genes function *in vivo*.

A truncated version of *dlx3* and a chimera consisting of the amino-terminal half of *dlx3* and the C-terminal half of *dlx2* were ubiquitously expressed in zebrafish embryos to determine the role of Dlx3 in the expression of *dlx4*, since *dlx3* was shown to be expressed in the presumptive otic placode (a thickening of the ectoderm that will eventually produce the inner ear (Ekker et al. 1992; Ellies et al. 1997 b)) a few hours before the expression of *dlx4*. Injection of mRNA has effectively been used previously to ubiquitously express protein products in zebrafish embryos (Kelly et al. 1995a, 1995b; de Vries et al. 1996).

Injection of full length *dlx2* and *dlx3* produced a small percentage of fish where *dlx4* expression was reduced or loss and the overall morphology were affected by those

injections (16%-17% respectively). It was shown previously by Zerucha et al who used a different expression vector (bluescript) for synthesising their mRNA, that theubiquitous expression, at the same concentrations, of full length *dlx3* was not disturbing at all the expression of *dlx4* (Zerucha et al., 1997). The percentage of dead embryos that they observed was approximately the same as the controls. In the present experiment, increasing stability of the injected mRNA, conferred by the CS2+ vector consequently increased the number of intact mRNA molecules. The higher concentrations of mRNA in the cells probably favour unspecific interactions with other endogenous proteins. This may have led to toxicity, thus creating morphological abnormality and higher incidence of dead embryos.

The injection of truncated *dlx2* produced a loss of or reduction in *dlx4* expression in the ears, visceral arches and the forebrain in 19% of the embryos injected. In addition, approximately 30% of them were morphologically affected. These results can not be compared with those of Zerucha et al, since they have not injected such construct. Does truncated *dlx2* affect specifically *dlx4* expression? It seems not since injections of full length *dlx2* and *dlx3* affected also embryos in 16% and 17% of them respectively and were shown by Zerucha et al not disturbing *dlx4* expression. Also, Dlx2 protein has not been shown yet to be a regulator of *dlx4* or *dlx3* expression. This perturbation may be due, as for full length *dlx2* and *dlx3*, to an increased stability of the mRNA conferred by CS2+ vector, which in turn would engender toxicity. It has been shown by Zerucha et al. that injection of lacZ mRNA was not disturbing the development. Thus the toxic effect observed in the present work would not be due to simply inject mRNA. It maybe due though to unspecific binding of the injected mRNA with endogenous proteins, which in turn would affect their normal function.

Ubiquitous expression of truncated *dlx3* failed to produce a significant percentage of embryos presenting phenotypes. *Dlx4* expression was affected in only 5% of the embryos injected and their morphology was normal. However, Zerucha et al demonstrated that ubiquitous expression of the same truncated Dlx3 version did affect *dlx4* expression. They showed that injection of the same concentration of a mRNA solution coding for truncated Dlx3 was perturbing *dlx4* expression in the ears, the visceral arches and the forebrain in 16% of the embryos injected. We used the expression vector CS2+ instead of bluescript to obtain greater percentages of affected embryos, but, in this case, a lower percentage of affected embryos was observed. The reason for this paradoxical result is, at present, unclear.

The injection of the chimera coding for the amino-terminal end of Dlx3 and the homeodomain and the carboxy-terminal end of Dlx2 produced various phenotypes including severe malformations in approximately 35% of the embryos. The different phenotypes related to *dlx4* expression are described in table 3-3. The phenotypes mostly seen are loss or reduction of *dlx4* expression in the arches only (12%), in the ears, the forebrain and the visceral arches (13%) and in the visceral arches and the ears (11%). Zerucha et al obtained a total of 24% affected embryos injected with the chimera where 12.5% of them had a loss or reduction of dlx4 expression in the visceral arches only and another 12.5% showed a loss or a reduction in the ears, visceral arches and forebrain. No morphological abnormalities were observed nor any increase in embryonic lethality. The use of the CS2+ vector in this present work produced higher percentages of embryos where *dlx4* expression is loss or reduced. The greatest effects on *dlx4* expression in this present work, were observed in the otic vesicle and in the visceral arches. The loss or reductions of *dlx4* expression in the otic vesicle after injection of the Dlx3-Dlx2 chimera suggested that Dlx3 was involved in the regulation of other *dlx* genes subsequently expressed in the same cells of the otic vesicle, such as *dlx4* and *dlx6*. This hypothesis is

supported by a co-transfection experiment where it was demonstrated that Dlx3 is able to activate transcription through a DNA fragment from the immediate *dlx4*-5'- flanking region (Zerucha et al. 1997)

A surprising phenotype observed was the one where the expression of *dlx4* is reduced in the ventral forebrain. It is difficult to explain this result since it was shown previously that Dlx3 is not expressed in forebrain cells (Akimenko et al. 1994). Therefore, it is not possible to explain this observation in terms of interference with endogenous Dlx3 function. It is possible though that the chimera disturbed the function of another *dlx* gene expressed in those cells, like *dlx2*.

Despite the similarities between Dlx proteins, specifically Dlx2 and Dlx3, in the homeodomain (56 identical amino acids out of 61) as well as in the 18 amino acid region, we have observed that a chimera consisting of the amino-terminal half of Dlx3 and the carboxy-terminal half of Dlx2 greatly interfered with the regulation of *dlx4* expression while overexpression of the entire Dlx3 had little effect on *dlx4*. Two mechanisms are proposed to explain these results. The first mechanism would imply that the homeodomains of Dlx2 and of Dlx3 do not show the same DNA binding specificities. The Dlx2 homeodomain is not completely identical to the Dlx3 homeodomain: four of the five differences between the Dlx3 and Dlx2 homeodomains are found downstream of the XhoI site used to create the chimeric cDNA. It is possible that these four amino acids influence the DNA binding specificities of the two proteins. Also it is important to take into account that the amino acids located downstream of the Dlx2 homeodomain may influence the DNA binding specificity and target the Dlx3-Dlx2 chimera to different sites in the genome to the one preferred by Dlx3. Therefore we can imagine that the chimera cannot bind to the Dlx3 target sites, including, sequences in the dlx4 5' flanking region that are activated by Dlx3 in cell culture. Thus, we can predict that Dlx3-Dlx2 chimera

sequesters a factor that interacts with the amino-terminal region of Dlx3. Supposing that this factor is essential for *dlx4* expression, its sequestration would result in reduction of *dlx4* expression since it is less available for endogenous Dlx3 (figure 4-1 II).

The second mechanism proposed that the homeodomain of the Dlx3-Dlx2 chimera binds DNA with the same specificity as Dlx3. We can suppose that the Dlx3-Dlx2 chimera competes with Dlx3 for binding sites on the DNA but cannot make the proper protein-protein interaction necessary for Dlx4 activation (figure4-1 II). This model implies that Dlx3 has an important functional domain downstream of its homeodomain that is not shared by Dlx2. This unknown domain would confer specificity to the protein.

4.3. Craniofacial dysmorphogenesis after injection of a chimera Dlx3-Dlx2

As a mentioned above, the greatest effects on *dlx4* expression seen when the chimera was ubiquitously expressed in zebrafish embryos, are in the otic vesicle and in the visceral arches. That is why we decided to study the effect of the overexpression of the chimera on the craniofacial cartilage structures and the inner ear formation when the fish are 5 days old and these structures have undergone a large part of their developmental program.

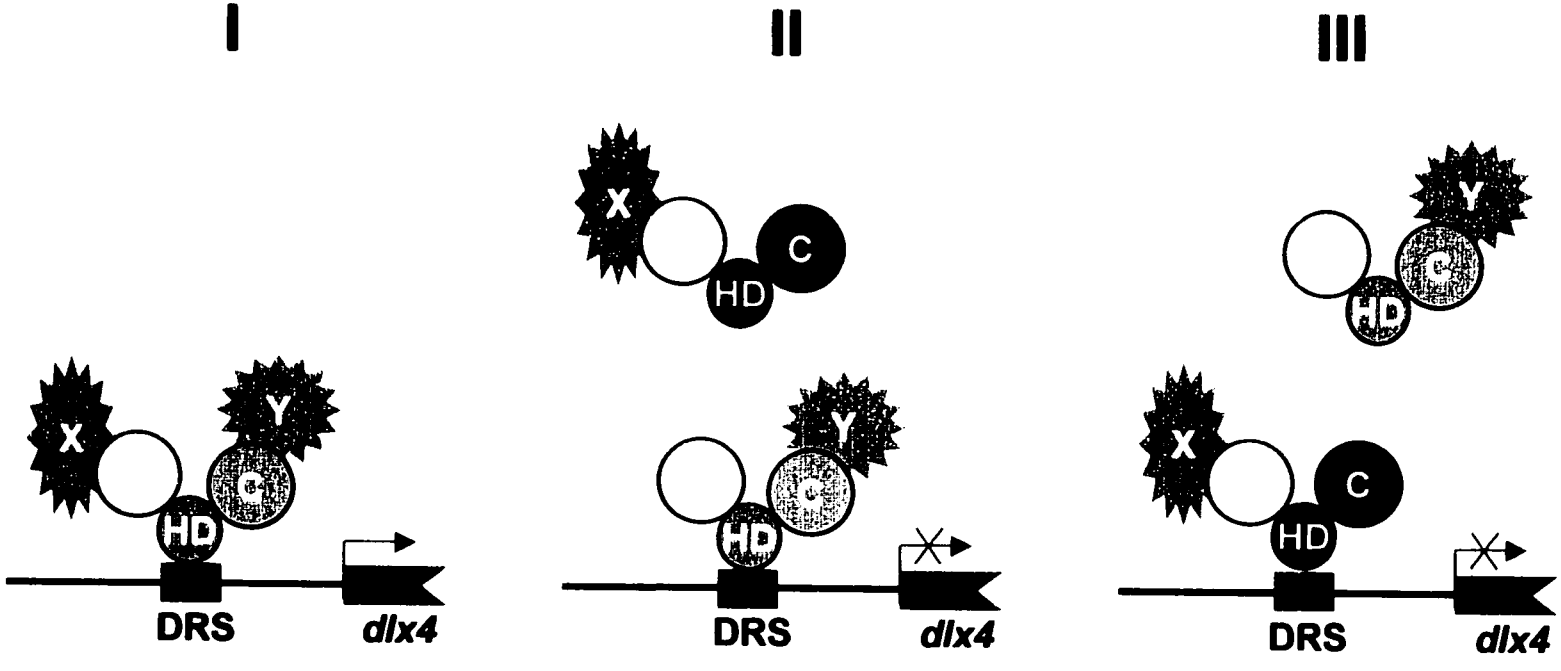
First the cartilage of 5 days post-fertilisation larvae were stained with alcian blue. This staining allowed us to examine the morphology of the cartilage. Six percent of the fish that we examined presented malformation of certain cartilage, not all of them. The Meckel's cartilage, the ceratohyal cartilage, the basihyal cartilage and the trabeculae. The Meckels cartilage shape was larger compared to controls (figure 3-10 B-D-F) and in some of them its normal U shape was not well formed (figure 3-10 C-E). The ceratohyal

cartilage did not present the characteristical V shape, as shown in control larvae (figure 3-10 C-E-F). Also, in some of the larvae, its morphogenesis appeared incomplete (figure 3-10 B-D). In some of the fishes, the basihyal was not detectable (figure 3-10 B-C) whereas in other ones, it was missplaced e.g. not at the junction of the ceratohyal (figure 3-10 D). The trabeculae which normally have a V shape, seemed to be fused together as shown in figure 3-10 B-F). The loss or reduction of *dlx4* expression in the visceral arches could have led to an impairment in the morphogenesis of the cartilage. The different cartilage elements affected in this work, originate from two distinct regions of the visceral arches that are called mandibular arch and hyoid arch. In each region a few members of the dlx gene family are expressed. The expression of *dlx4* is detected in all of them. It seems that the impairment of *dlx4* expression, as the result of chimera mRNA injections, can cause malformations of the cartilages but not their complete loss although the evidence is, at this stage only correlative.

It was reported previously that a loss of *dlx* gene expression correlated with a specific craniofacial cartilage dysmorphogenesis in retinoic acid treated zebrafish embryos. (Ellies et al., 1997b). In this study, it was observed that retinoic acid, a substance known to cause abnormal craniofacial cartilage development in other vertebrates, resulted in dose and stage dependent losses of *dlx* homeobox gene expression in several regions of the embryo. They also found that loss of *dlx* expression correlated either with the loss of cartilage elements originating from hindbrain neural crest cells or with abnormal morphology of these elements. Cartilage elements that originate from the midbrain neural crest cells such as the trabeculae, which do not express *dlx* genes, are less affected (Ellies, et al 1997b). From their results they suggested that *dlx* genes are an

Figure 4-1: Proposed model explaining the negative effects of the Dlx3-Dlx2 chimera on the activation of *dlx4* by the endogenous Dlx3.

(I) Wild-type transcriptional activation of *dlx4* by Dlx3. The *dlx4* transcription unit is represented by a dark orange box and transcription is indicated by the arrow. The Dlx3 protein is represented as three circles, each representing one protein domain: N, a putative domain in the N-terminal region; HD, the homeodomain; C, a putative domain in the carboxy-terminal region of the protein. Dlx3 is shown interacting via the homeodomain at its DNA recognition site (DRS) located in the *dlx4* 5'-flanking region. Also shown are two proposed yet unidentified cofactors interacting with the Dlx3 protein. These interactions would be required for proper activation of *dlx4* expression. (II) Squelching of cofactor X by Dlx3-Dlx2 chimera unable to bind DNA. This impairs endogenous Dlx3 function at its regulatory sites upstream of *dlx4* by competing for factor X. (III) The Dlx3-Dlx2 chimera is able to bind DNA and competes for the DRS with the endogenous Dlx3. The chimera would not be able to interact with factor Y because its Dlx2 C-terminal fragment, which is different from the Dlx3 C-terminal fragment. The binding of factor Y to C-terminal region of Dlx3 would be necessary for *dlx4* expression.


I
II
III
X
Y
C
HD
DRS
dlx4

important part of a multi-step process in the development of a subset of craniofacial cartilage elements (Ellies, et al 1997b). In the mouse the loss of *dlx* gene expression located in the visceral arches was also associated with craniofacial dysmorphogenesis. Selective inactivation of Dlx2 (Qui et al 1995), Dlx1, or both Dlx1 and Dlx2 (Qui et al. 1997) in the mouse results in malformation of craniofacial cartilage elements.

The correlation between loss of *dlx* expression and craniofacial abnormalities could also apply in this case. All the results suggest that *dlx4* might play an important role in craniofacial cartilage morphogenesis. Since there was no complete loss of these elements, it is possible that other *dlx* genes expressed in the visceral arches functionally compensate in part for the loss or reduced expression of *dlx4*.

Other morphological phenotypes were observed: an eye missing and curved tail which were surprising (figure 3-10 and 3-9). *Dlx4* expression is not thought to be involved in the eye development since no expression was observed in this region. However it is expressed in the ventral forebrain, which is nearby the eye formation site. The loss or reduction of its expression due to injections of chimera would maybe indirectly impair formation of an eye. The side-curved tail phenotype observed in this work as never been reported so far, only a tail curved upwards has been observed as a result of injections. The phenotype observed here is probably not specific to the loss of *dlx4* expression, since *dlx* expression as not been visualized in the body axis. It may be due to the injection of mRNA.

4.4. The inner ear morphology after injection of a chimera Dlx3-Dlx2.

Ubiquitous expressed of a chimera Dlx3-Dlx2 in zebrafish embryos seemed not to affect the morphogenesis of the inner ear of 5 days old embryos. All structures like the semicircular canals, the maculae, the cristae and the hair cells were present and seemed to be normal. Even fishes with large morphological defects did not show any structural abnormalities in the inner ear.

The first *dlx* genes expressed in the otic vesicle are *dlx3* and *dlx7*. Their patterns of expression are the same. Their expression is followed by *dlx4* and *dlx6* expression. The functional involvement of Dlx4 during patterning has not been reported so far, we only know it is expressed in the otic vesicle at around 14 hours post-fertilisation. Only Dlx3 has been studied more extensively. It is thought that Dlx3 is involve during the induction of the otic placode (Ekker et al. 1992), but it is not known to date what role it plays during patterning of the otic vesicle. It has been shown previously that injection of the Dlx3-Dlx2 chimera did affect *dlx4* expression (Zerucha et al 1997) in the otic vesicle. However, the patterning of the otic vesicle was not examined. In this study, we showed that a reduction or loss of *dlx4* expression did not influence the patterning of the inner ear. It seems that induction of the otic placode which involve *dlx3* expression is sufficient for proper patterning of the inner ear, whereas *dlx4* which is expressed after induction is not necessary.

The observations made on the morphology of the otic vesicle after injection of the Dlx3-Dlx2 chimera that affect *dlx4* expression suggested that *dlx4* is not absolutely necessary for normal development of the inner ear. It seems though, that Dlx3 is very important since it is implicated in the early events during the otic placode formation that is why Dlx3 is often used as a marker of the otic development (Mendosa et al., 1999).

5. Conclusions:

First, from the transfection experiments we were able to identify an activation domain located in the amino terminal region of Dlx1, Dlx2 and Dlx4 proteins. Also we were able to show that these proteins were activating transcription at different levels. Secondly, from the injection experiments, we can suggest that Dlx proteins play specific role during development. We showed that a chimera consisting in the amino terminal half of Dlx3 and the homeodomain and the carboxy terminal half of dlx2 cannot substitute for the function of Dlx3 protein, since perturbed *dlx4* expression was observed. The perturbed expression of *dlx4* after injection of the chimera suggests also that cross-regulation between Dlx proteins is a possible mechanism by which they exert their effect.

REFERENCES:

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

Akimenko, M. A., Ekker, M., Wegner, J., Lin, W., and Westerfield, M. (1994). Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head. J Neurosci **14**, 3475-86.

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R.K., Langeland, J., Prince, V., Wang, Y. L., Westerfield, M., Ekker, M., and Postlethwait, J. H. (1998). Zebrafish hox clusters and vertebrate genome evolution. Science **282**, 1711-4.

Anderson, S. A., Eisenstat, D. D., Shi, L., and Rubenstein, J. L. (1997b). Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes. Science **278**, 474-476.

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**, 265-272.

Asano, M., Emori, Y., Saigo, K., and Shiokawa, K. (1992). Isolation and characterization of a Xenopus cDNA which encodes a homeodomain highly homologous to Drosophila Distal-less. J Biol Chem **267**, 5044-7.

Beachy, P. A., Krasnow, M. A., Gavis, E. R., and Hogness, D. S. (1988). An Ultrabithorax protein binds sequences near its own and the *Antennapedia* P1 promoters. Cell **55**, 1069-81.

Beauchemin, M., and Savard, P. (1992). Two distal-less related homeobox-containing genes expressed in regeneration blastemas of the newt. Dev Biol **154**, 55-65.

Bodner, M., Castrillo, J. L., Theill, L. E., Deerinck, T., Ellisman, M., and Karin, M. (1988). The pituitary-specific transcription factor GFH-1 is a homeobox-containing protein. Cell **55**, 505-518.

Botfield, M. C., Janecso, A., and Weiss, M. A. (1992). Biochemical characterization of the Oct-2 POU domain with implication for bipartite DNA recognition. Biochemistry **31**, 5841-5848.

Boulikas, T., (1993). Homeodomain protein binding sites, inverted repeats, and nuclear matrix attachment regions along the human beta-globin gene complex. J. Cell. Biochem. **52**, 23-36.

Bourbon, H.-M., Martin-Blanco, E., Rosen, D., and Kornberg, T. B. (1995). Phosphorylation of the *Drosophila* Engrailed protein at a site outside its homeodomain enhance DNA binding. J. Biol. Chem. **270**, 11130-11139.

Brand, A. H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development **118**, 401-415.

Carroll, S. B., and Scott, M. P. (1985). Localization of the *fushi tarazu* protein during *Drosophila* embryogenesis. Cell **43**, 47-57.

Clerc, R. G., Corcoran, L. M., LeBowitz, J. H., Baltimore, D., and Sharp, P. A. (1988). The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. Genes & Dev. **2**, 1570-1581.

Cohen, S. M., Bronner, G., Kuttner, F., Jurgens, G., and Jackle, H. (1989). Distal-less encodes a homoeodomain protein required for limb development in Drosophila. Nature **338**, 432-4.

Cohen, S. M. (1990). Specification of limb development in the Drosophila embryo by positional cues from segmentation genes. Nature **343**, 173-7.

Desplan, C., Theis, J., and O'Farrel, P. H. (1985). The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity. Nature **318**, 630-5.

Desplan, C., Theis, J., and O'Farrel, P. H. (1988). The sequence specificity of homeodomain-DNA interaction. Cell **54**, 1081-90.

deVries, C. J. M., Boer, J. d., Joore, J., Strahle, U., Achterberg, T. A. E. v., Hylebroeck, D., Verschueren, K., Miyazono, K., Raajj, A. J. M. v. d. E.-v., and Zivkovic, D. (1996). Active Complex formation of type I and type II activin and TGF beta receptors in vivo as studied by overexpression in zebrafish embryos. Mech Dev *54*, 225-236.

Dirksen, M. L., Mathers, P., and Jamrich, M. (1993). Expression of a Xenopus Distal-less homeobox gene involved in forebrain and cranio-facial development. Mech Dev **41**, 121-128.

Dollé, P., Price, M., and Duboule, D. (1992). Expression of the murine Dlx-1 homeobox gene during facial, ocular and limb development. Differentiation **49**, 93-9.

Dong, J., Hung, L.-H., Strome, R., and Krause, M. (1998). A phosphorylation site in the Ftz homeodomain is required for activity. EMBO J. **17**, 2308-2318.

Ekker, M., Akimenko, M. A., Bremiller, R., and Westerfield, M. (1992). Regional expression of three homeobox transcripts in the inner ear of zebrafish embryos. Neuron **9**, 27-35.

Ellies, D. L., Stock, D. W., Hatch, G., Giroux, G., Weiss, K. M., and Ekker, M. (1997a). Relationship between the genomic organization and the overlapping embryonic expression patterns of the zebrafish dlx genes. Genomics **45**, 580-90.

Ellies, D. L., Langille, R. M., Martin, C. C., Akimenko, M. A., and Ekker, M. (1997b). Specific craniofacial cartilage dysmorphogenesis coincides with a loss of dlx gene expression in retinoic acid-treated zebrafish embryos. Mech Dev **61**, 23-36.

Faiella, A., Zappavigna, V., Mavillo, F., and Boncinelli, E. (1994). Inhibition of retinoic acid-induced activation of 3' human HOXB genes by antisense oligonucleotides affects sequential activation of genes located upstram in the four HOX clusters. Proc. Natl. Acad. Sci. U.S.A. **91**, 5335-5339.

Fainsod, A., Bogarad, L. D., Ruusala, T., Lubin, M., Crothers, D. M., and Ruddle, F. H. (1986). The homeodomain of a murine protein binds 5' to its own homeobox. Proc. Natl. Acad. Sci. U.S.A. **83**, 9532-6.

Feledy, J. A., Morasso, M. I., Jang, S. I., and Sargent, T. D. (1999). Transcriptional activation by the homeodomain protein Distal-less 3. Nucleic Acids Res. **27**, 764-770.

Ferrari, D., Sumoy, L., Gannon, J., Sun, H., Brown, A. M., Upholt, W. B., and Kosher, R. A. (1995). The expression pattern of the Distal-less homeobox-containing gene Dlx- 5 in the developing chick limb bud suggests its involvement in apical ectodermal ridge activity, pattern formation, and cartilage differentiation. Mech Dev **52**, 257-64.

Finney, M., Ruvkun, G., and Horvitz, H. R. (1988). The C. elegans cell lineage and differentiation gene unc-86 encodes a protein with a homeodomain and extended similarity to transcription factors. Cell **55** 757-769.

Gavis, E. R., and Hogness, D. S. (1991). Phosphorylation, expression, and function of the Ultrabithorax protein family in *Drosophila melanogaster*.

Gay, N. J., Poole, S. J., and Kornberg, T. B. (1988). The *Drosophila engrailed* protein is phosphorylated by a serine-specific protein kinase. Nucleic Acids Res. **16**, 6637-6647.

Gay, N. J., Poole, S. J., and Kornberg, T. B. (1988). Association of the *Drosophila melanogaster* engrailed protein with specific soluble nuclear protein complexes. EMBO J. 7, 4291-4297.

Gehring, W., (1994). The discovery of the homeobox. in Guidebook to the homeobox genes., Oxford University Press, New York.

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

Gehring, W. J., (1987). Homeoboxes in the study of development. Science **236**, 1245-1252.

Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984). Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Drosophila* embryonic development. Cell **37**, 833-841.

Han, K., Levine, M. S., and Manley, J. L. (1989). Synergistic activation and repression of transcription by *Drosophila* Homeobox proteins. Cell **56**, 573-583.

Hanes, S. D., and Brent, R. (1989). DNA specificity of the bicoid activator protein. Cell **57**, 1275-1283.

Hayashi, S., and Scott, M. P. (1990). What determines the specificity of action of *Drosophila* homeodomain proteins? Cell **63**, 883-894.

Herr, W., Sturm, R. A., Clerc, R. G., Corcoran, L. M., Baltimore, D., Sharp, P. A., Ingraham, H. A., Rosenfeld, M. G., Finney, M., Ruvkun, G., and Horvitz, H. R. (1988). The POU domain: A large conserved region in the mammalian pit-1, oct-1, oct-2, and *Caenorhabditis elegans unc-86* gene products. Genes & Dev. **2**, 1513-1516.

Hill, C. S., and Treisman, R. (1995). Transcriptional regulation by extracellular signals: mechanisms and specificity. Cell **80**, 199-211.

Hiromi, Y., Kuroiwa, A., and Gehring, W. J. (1985). Control elements of the *Drosophila* segmentation gene *fushi tarazu.* Cell **43**, 603-613.

Hoey, T., and Levine, M. (1988). Divergent homeobox proteins recognize similar DNA sequences in *Drosophila.* Nature **332**, 858-61.

Hunter, T., and Karin, M. (1992). The regulation of transcription by phosphorylation. Cell **70**, 375-387.

Ingraham, H. A., Flynn, S. E., Voss, J. W., Albert, V. R., Kapiloff, M. S., Wilson, L., and Rosenfeld, M. G. (1990). The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA binding and DNA-dependent Pit-1-Pit-1 interactions. Cell **61**, 1021-1033.

Jaffe, L., Ryoo, H., and Mann, R. S. (1997). A role for phosphorylation by casein kinase II in modulating *Antennapedia* activity in *Drosophila.* Genes & Dev. **11**, 1327-1340.

Jaynes, J. B., and O'Farrell, P. H. (1988). Activation and repression of transcription by homeodomain-containing proteins that bind a common site. Nature **336**, 744-9.

Kalionis, B., and O'Farrell, P. H. (1993). A universal target sequence is bound *in vitro* by diverse homeodomains. Mech. Dev. **43**, 57-70.

Karin, M. (1994). Signal transduction from the cell surface to the nucleus through the phosphorylation of transcription factors. Curr. Opin. Cell Biol. **6**, 415-424.

Kelly, G. M., Erezyilmas, D. F., and Moon, R. T. (1995a). Induction of a secondary embryonic axis in zebrafish occurs following the overexpression of b catenin. Mech Dev **53**, 261-273.

Kelly, G. M., Greenstein, P., Erezyilmas, D. F., and Moon, R. T. (1995b). Zebrafish wnt8 and wnt8b share a common activity but are involved in distinct developmental pathways. Development **121**, 1787-1799.

Kimmel, C.B. (1989). Genetics and early development of zebrafish. Trends in genetics,. **5**: p 283-288.

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

Klemm, J. D., and Pabo, C. O. (1996). Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage. Genes & Dev. **10**, 27-36.

Klemm, J. D., Rould, M. A., Aurora, R., Herr, W., and Pabo, C. O. (1994). Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. Cell **77**, 21-32.

Krasnow, M. A., Saffman, E. E., Kornfeld, K., and Hogness, D. S. (1989). Transcriptional activation and repression by Ultrabithorax proteins in cultured *Drosophila* cells. Cell **57**, 1031-43.

Krause, H. M., Klemenz, R., and Gehring, W. J. (1988). Expression, modification, and localization of the *fushi tarazu* protein in *Drosophila* embryos. Genes & Dev. **2**, 1021-1036.

Krause, H. M., and Gehring, W. J. (1989). Stage-specific phosphorylation of the *fushi tarazu* protein during *Drosophila* development. EMBO J. **8**, 1197-1204.

Kristie, T. M., and Sharp, P. A. (1990). Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV α-trans-activator protein. Genes & Dev. **4** 2383-96.

Laughon, A., and Scott, M. P. (1984). Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature **310**, 25-31.

Laughon, A., Howell, W., and Scott, M. P. (1988). The interaction of proteins encoded by *Drosophila* homeotic and segmentation genes with specific DNA sequences. Development **104**, 75-83.

Lu, Q., Knoepfler, P. S., Scheele, J., Wright, D. D., and Kamps, M. P. (1995). Both 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.

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

Mc Ginnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. Cell, **68**: p 283-302.

McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984). A conserved DNA sequence in homeotic genes of the *Drosophila Antennapedia* and *bithorax* complexes. Nature **308**, 428-33.

McGuinness, T., Porteus, M. H., Smiga, S., Bulfone, A., Kingsley, C., Qiu, M., Liu, J. K., Long, J. E., Xu, D., and Rubenstein, J. L. (1996). Sequence, organization, and transcription of the Dlx-1 and Dlx-2 locus. Genomics **35**, 473-85.

Mendonsa, E. S., and Riley, B. B. (1999). Genetic analysis of tissue interactions required for otic placode induction in the zebrafish. Developmental Biol. **206**, 100-112.

Mullen, L. M., Bryant, S. V., Torok, M. A., Blumberg, B., and Gardiner, D. M. (1996). Nerve dependency of regeneration: the role of Distal-less and FGF signaling in amphibian limb regeneration. Development **122**, 3487-97.

Muller, M., Affolter, M., Leupin, W., Otting, G., Wuthrich, K., and Gehring, W. J. (1988). Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. EMBO J **7**, 4299-304.

Nakamura, S., Stock, D. W., Wydner, K. L., Bollekens, J. A., Takeshita, K., Nagai, B. M., Chiba, S., Kitamura, T., Freeland, T. M., Zhao, Z., Minowada, J., Lawrence, J. B., Weiss, K. M., and Ruddle, F. H. (1996). Genomic analysis of a new mammalian distal-less gene: Dlx7. Genomics **38**, 314-24.

Otting, G., Qian, Y. Q., Billeter, M., Müller, M., Affolter, M., Gerhing, W. J., and Wüthrich, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. EMBO J. **9**, 3085-3092.

Ozcelik, T., Porteus, M. H., Rubenstein, J. L., and Francke, U. (1992). DLX2 (TES1), a homeobox gene of the Distal-less family, assigned to conserved regions on human and mouse chromosomes 2. Genomics **13**, 1157-61.

Pabo, C. O., and Sauer, R. T. (1984). Protein-DNA recognition. Annu. Rev. Biochem **53**, 293-321.

Papalopulu, N., and Kintner, C. (1993). Xenopus Distal-less related homeobox genes are expressed in the developing forebrain and are induced by planar signals. Development **117**, 961-75.

Perez, F., Joliot, A., Bloch, G. E., Zahraoui, A., Triller, A., and Prochiantz, A. (1992).Antennapedia homeobox as a signal for the cellular internalization and nuclear addressing of a small exogenous peptide. J. Cell. Sci. **102**, 717-722.

Phelan, M., Sadoul, R., and Featherstone, M. S. (1994), Functional differences between Hox proteins conferred by two residues in the homeodomain N-terminal arm. Mol. Cell. Biol. **14**, 5066-5075.

Pöpperl, H., Bienz, M., Studer, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S., and Krumlauf, R. (1995). Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon exd/pbx. Cell **81**, 1031-42.

Porteus, M. H., Bulfone, A., Ciaranello, R. D., and Rubenstein, J. L. (1991). Isolation and characterization of a novel cDNA clone encoding a homeodomain that is developmentally regulated in the ventral forebrain [published erratum appears in Neuron 1992 Jul;9(1):187]. Neuron **7**, 221-9.

Price, M., Lemaistre, M., Pischetola, M., Di Lauro, R., and Duboule, D. (1991). A mouse gene related to Distal-less shows a restricted expression in the developing forebrain. Nature **351**, 748-51.

Qiu, M., Bulfone, A., Martinez, S., Meneses, J. J., Shimamura, K., Pedersen, R. A., and Rubenstein, J. L. (1995). Null mutation of Dlx-2 results in abnormal morphogenesis of proximal first and second branchial arch derivatives and abnormal differentiation in the forebrain. Genes Dev **9**, 2523-38.

Qiu, M., Bulfone, A., Ghattas, I., Meneses, J. J., Christensen, L., Sharpe, P. T., Presley, R., Pedersen, R. A., and Rubenstein, J. L. (1997). Role of the Dlx homeobox genes in proximodistal patterning of the branchial arches: mutations of Dlx-1, Dlx-2, and Dlx-1 and -2 alter morphogenesis of proximal skeletal and soft tissue structures derived from the first and second arches. Dev Biol **185**, 165-84.

Quian, Y. Q., Billleter, M., Otting, G., Müller, M., Gerhing, W. J., and Wüthrich, K. (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with procaryotic repressors. Cell **59**, 573-80.

Rauskolb, C., Peifer, M., and Wieschaus, E. (1993). Extradenticle, a regulator of homeotic gene activity, is a homologue of the homeobox-containing human proto-oncogene Pbx1. Cell **74**, 1101-1112.

Robinson, G. W., Wray, S., and Mahon, K. A. (1991). Spatially restricted expression of a member of a new family of murine Distal-less homeobox genes in the developing forebrain. New Biol **3**, 1183-94.

Ronchi, E., Treisman, J., Dostatni, N., Struhl, G., and Desplan, C. (1993). Down-regulation of the *Drosophila* morphogen *bicoid* by the *torso* receptor-mediated signal transduction cascade. Cell **74**, 347-355.

Ruddle, F. H., Bartels, J. L., Bentley, K. L., Kappen, C., Murtha, M. T., and Pendelton, J. W. (1994a). Evolution of Hox genes. Annu. Rev. Genet. **28**, 423-442.

Ruddle, F. H., Bentley, K. L., Murtha, M. T., and Risch, N. (1994b). Gene loss and gain in the evolution of the vertebrates. Development Suppl. 155-161.

Ryoo, H. M., Hoffmann, H. M., Beumer, T., Frenkel, B., Towler, D. A., Stein, G. S., Stein, J. L., van Wijnen, A. J., and Lian, J. B. (1997). Stage-specific expression of Dlx-5 during osteoblast differentiation: involvement in regulation of osteocalcin gene expression. Mol Endocrinol **11**, 1681-94.

Scherer, S. W., Heng, H. H., Robinson, G. W., Mahon, K. A., Evans, J. P., and Tsui, L. C. (1995). Assignment of the human homolog of mouse Dlx3 to chromosome 17q21.3-q22 by analysis of somatic cell hybrids and fluorescence in situ hybridization. Mamm Genome **6**, 310-1.

Schier, A. F., and Gehring, W. J. (1993). Functional specificity of the homeodomain protein fushi tarazu: the role of DNA binding specificity in vivo. Proc. Natl. Acad. Sci. U.S.A. **90**, 1450-1454.

Scott, M. P., and Weiner, A. J. (1984). Structural relationships among genes that control development: sequence homology between the *Antennapedia, Ulthrabithorax* and *fushi tarazu* loci of *Drosophila.* Proc. Natl. Acad. Sci. U.S.A. **81**, 4115-4119.

Scott, M. P., Tamkun, J. W., and Hartzell, G. W. (1989). The structure and function of the homeodomain. Biochim. Biophys. Acta **989**, 25-48.

Scott, M. P., Tamkun, J. W., and Hartzell, G. W. (1989). Biochim. Biophys. Acta **989**, 25-48

Shepherd, J. C., McGinnis, W., Carrasco, A. E., De Robertis, E. M., and Gehring, W. J. (1984). Fly and frog homeodomains show homologies with yeast mating type regulatory proteins. Nature **310**, 70-71.

Shirasawa, T., Sakamoto, K., and Tkahashi, H. (1994). Molecular cloning and evolutional analysis of a mammalian homologue of the Distal-less 3 (Dlx-3) homeobox gene. FEBS Lett **351**, 380-4.

Simeone, A., Acampora, D., Pannese, M., D'Esposito, M., Stornaiuolo, A., Gulisano, M., Mallamaci, A., Kastury, K., Druck, T., Huebner, K., and et al. (1994). Cloning and characterization of two members of the vertebrate Dlx gene family. Proc Natl Acad Sci U S A **91**, 2250-4.

Stock, D. W., Ellies, D. L., Zhao, Z., Ekker, M., Ruddle, F. H., and Weiss, K. M. (1996). The evolution of the vertebrate Dlx gene family. Proc Natl Acad Sci U S A **93**, 10858-63.

Stuart, G. W., McMurray, J. V., and Westerfield, M. (1988). Replication, intergration, and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. Development **103**, 403-412.

Sturm, R. A., and Herr, W. (1988). The POU domain is a bipartite DNA-binding structure. Nature **336**, 601-604.

Vachon, G., Cohen, B., Pfeifle, C., McGuffin, M. E., Botas, J., and Cohen, S. M. (1992). Homeotic genes of the Bithorax complex repress limb development in the abdomen of the Drosophila embryo through the target gene Distal-less. Cell **71**, 437-50.

Verrijzer, C. P., Alkema, M. J., Van Weperen, W. W., Van Leeuwen, H. C., Starting, M. J., and van der Vliet, P. C. (1992). The DNA binding specificity of the bipartite POU domain and its subdomains. EMBO J. **11**, 4993-5003.

Viganò, M. A., Di Rocco, G., Zappavigna, V., and 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.

Wagner, M. J., Sharp, J. A., and Summers, W. C. (1981). Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type1. Proc. Natl. Acad. Sci. U.S.A. **78**, 1441-1445.

Westerfield, M. (1993). The Zebrafish Book., Eugene, OR: University of Oregon Press.

Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D., and Pabo, C. O. (1991). Crystal structure of a Matα2 homeodomain-operator complexe suggests a general model for homeodomain-DNA interactions. Cell **67**, 517-528.

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.

Zerucha, T., Muller, J. P., Chartrand, N., and Ekker, M. (1997). Cross-interactions between two members of the Dlx family of homeobox- containing genes during zebrafish development. Biochem Cell Biol **75**, 613-22.

Zerucha, T., Stühmer, T., Park, B. K., Long, Q., Yu, G., Hatch, G., Gambarotta, A., Schultz, J., Rubenstein, J. L. R., and Ekker, M. (1999). A Highly Conserved Enhancer In The Dlx5/Dlx6 Intergenic Region Is The Site Of Cross-Regulatory Interactions Between Dlx Genes In The Embryonic Forebrain. Submitted to Development March 1999.

Zhang, H., Hu, G., Wang, H., Sciavolino, P., Iler, N., Shen, M. M., and Abate-Shen, C. (1997). Heterodimerization of Msx and Dlx homeoproteins results in functional antagonism. Mol Cell Biol **17**, 2920-32.

Zhao, G. Q., Zhao, S., Zhou, X., Eberspaecher, H., Solursh, M., and de Crombrugghe, B. (1994). rDlx, a novel distal-less-like homeoprotein is expressed in developing cartilages and discrete neuronal tissues. Dev Biol **164**, 37-51.

Zwilling, S., Dieckmann, A., Pfisterer, P., Angel, P., and Wirth, T. (1997). Inducible expression and phosphorylation of coactivator BOB.1/OBF.1 in T cells. Science **277**, 221-225.