NOTE TO USERS

Initial Characterization of the CDP2/Cux2 Protein

Hélène Gingras

Biochemistry Department

McGill University, Montreal

October 2003

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

Your file Votre référence
ISBN: 0-612-98643-8
Our file Notre référence
ISBN: 0-612-98643-8

Table of Contents

Abstract

The *CDP1* (CCAAT-displacement protein1) and *cux1* (Cut homeobox1) genes were originally identified as the human and mouse orthologs of *Drosophila melanogaster cut*. CDP1/Cux1 is a transcription factor that has four DNA binding domains, three Cut repeat (CR) and one Cut homeodomain (HD) and can function both as repressor and activator. CDP1/Cux1 was shown to be modulated primarily by a variety of post-translational modifications including phosphorylation/dephosphorylation, specific proteolytic cleavage and acetylation. More recently, mammals were found to possess a second *cut* orthologs that was generated by gene duplication: *CDP2* in human, and *cux2* in mouse. *Cux2* mRNA expression was detected in various tissues during development in mouse and chicken and was found to be restricted to the neuronal system in the adult mouse. Nothing has been reported regarding the CDP2/Cux2 protein.

To begin to characterize the CDP2/Cux2 protein, I generated a number of reagents to analyze its expression and activity. Polyclonal antibodies were obtained by injecting rabbits with glutathione-S-transferase fusion proteins containing antigenic regions of the Cux2 protein. The antibodies were then characterized in Western blot analysis and electrophoretic mobility shift assays (EMSA). In total, 5 antibodies were produced against different regions of Cux2. These antibodies were able to specifically recognize CDP2 and Cux2 but not CDP1 or Cux1. Expression of the Cux2 protein was found in only one among 19 neuronal cell lines: the SH-SY5Y human neuroblastoma cell line. Histidine-tagged fusion protein containing various combinations of Cut repeats and Cut homeodomain were generated to investigate the DNA binding properties of CDP2/Cux2. The CR1CR2, CR2CR3HD and CR3HD domains were found to exhibit similar DNA binding specificities as the corresponding domains of CDP1/Cux2, however, analysis of DNA binding kinetics revealed that all of these combinations of domains made rapid but transient interactions with DNA. Mammalian expression vectors were engineered with epitope tags at the N- and C-termini of CDP2/Cux2. The full-length protein was found to localize in the nucleus and also to make a rapid but transient interaction with DNA. In contrast to CDP1/Cux1, the CDP2/Cux2 protein did not appear to be subject to specific proteolytic processing.

Résumé

Les gènes *CDP1* (CCAAT-displacement protein1) et *cux1* (Cut homeobox1) ont originalement été identifiés comme étant les orthologues humains et murins du gène *cut* de *Drosophila melanogaster*. CDP1/Cux1 est un facteur de transcription qui contient quatre domaines de liaison à l'ADN, trois "Cut repeats" (CR) et un homéodomaine Cut (HD), et peut agir comme répresseur et activateur. Les études ont montré que la régulation de CDP1/Cux1 se faisait principalement par des modifications post-traductionnelles incluant la phosphorylation/dephosphorylation, le processing protéolytique et l'acétylation. Récemment, un deuxième orthologue de *cut* a été identifié chez les mammifères: *CDP2* chez l'humain et *cux2* chez la souris. L'expression de l'ARNm de *cux2* fut détectée dans divers tissus lors du développement de la souris ou du poulet, mais seulement dans le système nerveux dans la souris adulte. On ne sait rien de la protéine CDP2/Cux2.

J'ai généré plusieurs réactifs afin d'étudier l'expression et l'activité de la protéine CDP2/Cux2. Des anticorps polyclonaux de lapin ont été produits à l'aide de protéines de fusion contenant la "glutathione-S-transferase" ainsi que diverses régions antigéniques de Cux2. J'ai caractérisé les anticorps dans des expériences de buvardage Western et des tests de mobilité électrophorétique (EMSA). Au total, j'ai produit cinq anticorps capables de reconnaître spécifiquement les protéines CDP2 et Cux2, mais non les protéines CDP1 et Cux1. Sur 19 lignées de cellules neuronales, seule la lignée de neuroblastome human SH-SY5Y exprimait la protéine CDP2. J'ai produit trois protéines de fusion avec une étiquette "histidine" et différentes combinaisons de domaines de liaison à l'ADN. Les domaines CR1CR2, CR2CR3HD et CR3HD ont montré les mêmes spécificité de liaison à l'ADN que les domaines correspondants des protéines CDP1/Cux1. Contrairement à CDP1/Cux1 cependant, l'interaction de tous ces domaines avec l'ADN était rapide et instable. J'ai produit des vecteurs d'expression pour cellules de mammifères avec des épitopes antigéniques aux extrémités amino- et carboxy-terminales des protéines CDP2 et Cux2. Les protéines entières se sont localisées dans le noyau et formaient une interaction transitoire avec l'ADN. Au contraire de CDP1/Cux1, les protéines CDP2 et Cux2 ne semblent pas être sujètes au clivage protéolytique spécifique.

Acknowledgment

In the beginning, I would like to thank my supervisor, Dr. Alain Nepveu, for his encouragement and supervision during my time in his laboratory. I would like to express my appreciation for his patience and guidance throughout my scientific development.

My appreciation to all the current members of Dr. Nepveu's laboratory: Ginette Bérubé, Mélanie Désilets, Brigitte Goulet, Dr. Ryoko Harada, Anie Monast, Lam LeDuy, Laurent Sansregret, Mary Truscott, Dr. Brian Wilson, Marianne Santaguida, Jonathan Vincent-Heroux, Yelena Markovic and Chantal Cadieux.

Thanks to Mary Truscott, Brigitte Goulet and Dr. Alain Nepveu for proofreading of the manuscript.

I would like also to thank Maria Krassilnikova, a former summer student who started the project before I arrived in the lab. She did a great job; her results are presented in Figure 1 and 7A.

The Cux2 constructs utilized in this project were prepared by Ginette Bérubé and Mélanie Désilets. I would like to acknowledge Lam LeDuy for his technical assistance and friendly support. A special thanks to my bay-mate, Mary Truscott, for her advice and support.

I am grateful to all the providers of neuronal cell lines: Dr. Rima Rozen, Dr. David Kaplan, Dr. Nathalie Lamarche, Dr. Xiang-Jiao Yang, Dr. Alan Peterson and Dr. Christian Sirard. I would like to thank Dr. Edith Hamel and Dr. Peter McPherson for the human and rat brain samples. A special attention to our Cux2 antibodies collaborator, Dr Oliver Cases, his work is very important for the determination of Cux2 localization in mice and human.

Last but not least, I would like to express my appreciation to my family, my friends and my life partner, you were all there behind me giving me constant encouragement and support throughout the last year, you are all precious to me.

List of Figures and tables **Page**

Introduction

1. *Cut* superclass of homeobox genes

Homeobox genes encode transcription factors that are involved in developmental decisions at many levels in a large range of organisms including yeast, insects, plants and mammals (21, 34). The homeobox is a 180-base pairs nucleotide sequence that encodes a DNA-binding motif called the homeodomain (109). Homeobox genes were first identified in the Antennapedia/Bithorax complexes of *Drosophila melanogaster* but were later found in higher metazoans. Early classification divided homeobox genes into the *Antennapedia* (*Antp*)-type genes and the diverged (non-*Antp*-type) genes (34, 64). Homeobox genes can now be classified into many classes on the basis of several criteria including organization within a gene cluster, sequence identity within the homeodomain and the presence of an additional DNA binding domain (19, 20, 104). The *cut* gene from *Drosophila melanogaster* gave its name to one of the classes because it is its oldest member (13). The *cut* superclass of homeobox genes has been divided into three classes: *CUX*, *ONECUT* and *SATB* (21) (refer to Appendix, Fig. A). While the Drosophila *cut*, the human *CDP1* and the mouse *cux1* genes contain three Cut repeats upstream of the homeobox in each species, there is also a *ONECUT* gene containing a single Cut repeat (12, 66, 72, 92, 126). SATB1 includes two Cut repeat-like domains and a divergent Cut-like homeodomain (31). The gene structure of *Drosophila* Cut, with three cut repeats and one homeodomain was subsequently shown to be conserved in vertebrate homologues (3, 92). Consequently, the Cut homologues were found to be encoded by orthologous genes that were derived from speciation events (100). Furthermore, in vertebrates, there are two *cut* homologues named CDP1 and CDP2 in humans, and Cux1 and Cux2 in mice. These were characterized to be paralogous genes that were derived from a gene duplication event (100).

2. Gene duplication

A gene can be duplicated in two ways, by cis-duplication or by trans-duplication. Cis-duplication events of genes occur early in animal evolution whereas a trans-duplication event occurs later. This will yield gene clusters on different chromosomes (24). New evidence has confirmed that gene duplication is the mechanism by which vertebrates

acquired more genes than their closest invertebrate relatives (54). This conclusion is based on the study of amphioxus and tunicates that are more appropriate animals for comparison to vertebrates than *Drosophila* and *Caenorhabditis elegans* (54, 84). For example, in agreement with the hypothetical model of genome duplication comparison of clustered homeobox genes, where the loss of a gene can be identified, suggests that one or two duplication events occurred, producing four clusters of Hox genes (55, 79, 96, 99, 108, 110, 115).

Three alternative models have been raised to describe the mechanism that preserves a large proportion of duplicated genes for long periods. The classical model, first stated by Ohno (1970), proposes that one copy remains normal and the other may be simply silenced due to degenerative mutations (Nonfunctionalisation; Degeneration). Another model proposes that one copy may acquire a new beneficial function and may become preserved by natural selection, while the other copy retains the original function (Neofunctionalisation). An additional model introduced a new conceptual framework, the duplication-degeneration-complementation (DDC) model (40) (refer to Appendix, Fig. B), in which gene preservation is explained by the fixation of degenerative mutations rather than by the fixation of new beneficial mutations. This model predicts that (1) degenerative mutations in regulatory elements can increase rather than reduce the probability of duplicate gene preservation and (2) the usual mechanism of duplicate gene preservation is the partitioning of ancestral functions rather than the evolution of new functions (40).

3. CCAAT displacement protein CDP1/Cux1

Transcriptional regulation of all genes is initiated by the specific binding of regulatory proteins called transcription factors to specific sites on DNA called promoter regions. Transcription factors employ a variety of mechanisms to recognize their DNA target sites. DNA recognition by a particular regulatory protein can be a complex multistep process (9, for a review see 133). However, we can simply describe it by a chemical process, in which the rate of the reaction is limited by the rate that the two components (i.e., the protein and the DNA site) are brought together via diffusion. Once brought together into a protein–DNA complex, the dissociation rate depends on the affinity, or chemical complementarity, between them. In general, the higher the number of favorable chemical

contacts between them the lower the dissociation rate and higher the affinity (9). Transcription factors contain at least one DNA binding domain that enables them to bind to DNA. Multiple DNA binding motifs have been identified: helix-turn-helix (HTH), zinc finger, helixloop-helix and leucine zipper. A common feature of the conserved DNA binding motifs is that they are shared by numerous transcription factors that present disparate biological processes (102). CDP1/Cux1 (CCAAT-displacement protein/cut homeobox) belongs to a family of transcription factors present in all metazoans, and is involved in the control of proliferation and differentiation (reviewed in 91). This family constitutes a unique group of homeoproteins that contain a Cut homeodomain as well as one or more Cut repeat DNA binding domains (91). The founding member of this family has been named Cut and is found in *Drosophila melanogaster*. In *Drosophila* embryos, *cut* expression was observed in the peripheral and central nervous system, malpighian tubules, tracheal histoblasts, cells surrounding the spiracles and adepithelial cells. In the adult, *cut* expression was observed in the same tissues and also in others (15). In *Drosophila melanogaster*, ablation of the *cut* gene caused embryonic lethality. In addition, a large number of phenotypes resulted from the insertion of transposable insulator sequences that interfered with the function of tissue-specific enhancers (23, 59, 60, 61, 87) (refer to Appendix, Fig. C). The affected tissues included the wings ("cut wing"), legs, external sense organs, Malpighian tubules, tracheal system and some structures in the central nervous systems (11, 16, 18, 58, 60, 76, 77). Genetic studies in *Drosophila melanogaster* indicated that Cut plays an important role in determining cell-type specificity in several tissues (13, 14, 16, 23, 33, 59, 62, 76, 77, 87). This was most thoroughly demonstrated in the peripheral nervous system (14, 16, 77, 78). For example, embryonic lethal Cut mutations caused the transformation of external sensory organs into internal (chordotonal) sensory organs (16). In contrast, when Cut expression was artificially elevated in embryos, all precursor cells gave rise to external sensory organs (14, 78). In addition, when there was no expression of Cut in tissues such as the wing and leg margin, it resulted in the truncation of the wing and a kinked femur. Cut was also shown to interact genetically with developmentally important genes like *notch* and *wingless* (4, 29, 32, 58, 85, 93). Interestingly, the dendritic arborization (da) sensory neuron in the embryonic and larval development of Drosophila shows distinct patterns of terminal dendrites branching whether

the expression of Cut is high or low (46). Loss of Cut reduced dendrite growth and class-specific terminal branching, whereas overexpression of Cut or a mammalian homolog in lower-level neurons resulted in transformations toward the branch morphology of high-Cut neurons. Thus, different levels of a homeoprotein can regulate distinct patterns of dendrite branching (46).

In mammals, the *cut* homologue was named CCAAT displacement protein (CDP1) because it was discovered as a protein that is capable of repressing transcription by competing with transcriptional activators for binding to the CCAAT elements of a gene promoter (7, 92).

Cut and CDP1 contain five evolutionarily conserved domains beginning at the amino terminus by a coiled-coil region, assumed to be implicated in protein-protein interactions, three Cut repeats and a homeodomain, which are involved in the interaction of the protein with DNA. In addition, the C-terminal region is not really conserved, however, it is rich in proline and alanine residues and active repression domains were identified within this region of the CDP1 protein. The full-length protein, CDP1/Cux1 p200, is proteolytically processed at the G1/S transition of the cell cycle, thereby generating the CDP1/Cux1 p110 isoform that contains three DNA binding domains, CR2, CR3, and HD (89). In contrast to p200, the p110 isoform is capable of stable interaction with DNA. Moreover, it was recently found that Cathepsin L could be involved in the processing of CDP1/Cux1 (Goulet et al. Unpublished data). Also, it was found that the p110 isoform is expressed at a higher level in uterine leiomyomas and in many transformed cell lines (90). Another short isoform, p75, was identified, and was shown to result from the initiation of transcription within intron 20 (44). This isoform displays slightly higher DNA binding stability than the p110 isoform, and its expression is restricted to certain tissues or cells including the thymus, T cells and the placenta. However, its expression was found to be aberrantly activated in breast tumor cell lines and primary human breast tumors (44). Those results, together with the finding that the p110 isoform is expressed at a higher level in uterine leiomyomas, suggest that alternative mechanisms may be selected in cancer cells to favor the expression of short CDP1/Cux1 isoforms.

Results from studies in mammals suggested that CDP1/Cux1 expression or activity might be restricted to proliferating cells (91). As a model, the expression of the mouse

Cux1 protein was inversely related to the degree of cellular differentiation in the kidney (52). Related to that, CDP1/Cux1 was reported to repress a large number of genes in precursor cells, in particular those genes expressed following terminal differentiation (7, 27, 35, 53, 62, 65, 67, 68, 74, 80, 82, 94, 95, 97, 112, 116, 120, 128, 137). The protein operates as a transcriptional repressor by at least two mechanisms: competition for binding site occupancy and active repression (82). The first mechanism involves the CCAAT displacement activity of the protein whereas the carboxy-terminal domain (CTD) of the protein that contains two active repression domains mediates active repression. Cell cycle regulators are also transcriptionally regulated by CDP/Cux1. CDP1/Cux1 can repress a reporter plasmid carrying the promoter of the p21Waf1/Cip1 gene (27). Furthermore, expression of another cyclin kinase inhibitor, p27, was down-regulated in CMV/Cux1 transgenic mice (69).

CDP1/Cux1 may also be involved in gene activation. A number of groups have identified binding sites for CDP1/Cux1 in the regulatory sequences of genes encoding histones and thymidine kinase (7, 36, 65,130). Additionally, CDP1/Cux1 was found to be part of the promoter complex HiNF-D, which is thought to be contributing to the transcriptional induction of some histone genes at the G1/S phase transition of the cell cycle (6, 129, 130). Recently, our laboratory has reported that CDP1/Cux1 p110, but not CDP/Cux p200, was capable of stimulating the expression of a reporter containing the promoter of the DNA polymerase alpha (DNA pol α) gene, which is up-regulated at the G1/S transition of the cell cycle and is needed for DNA synthesis (123). Moreover, expression of the endogenous DNA pol α gene was stimulated upon infection of cells with a retroviral vector expressing p110.

Cut proteins are unique in that they contain four DNA binding domains: the Cut homeodomain (HD) and the three Cut repeats (CR1, CR2 and CR3). The high degree of conservation of Cut repeats suggested that they might have an important biochemical function (12, 92). Indeed, Cut repeats were found to function as specific DNA binding domains (2, 5, 50, 51). Using histidine-tagged fusion proteins, various combinations of domains were found to bind to DNA with distinct affinities and kinetics: CR1CR2, CR3HD, CR1HD, CR2HD and CR2CR3HD (88). Interestingly, CR1CR2 displayed rapid on- and off-rates, and bound preferably to two CA/GAT sites organized as direct or inverted repeats

and separated by a variable number of nucleotides (88). In contrast, the CR2CR3HD and CR3HD bound more slowly, but stably to DNA and had preference for sequences conforming to the ATNNAT consensus (88). A histidine-tagged full length CDP1/Cux1 protein purified from insect cells exhibited DNA binding kinetics similar to that of CR1CR2, suggesting that the Cut homeodomain may not be active in the context of the full-length protein. Moreover, as expected, the CDP1/Cux1 p110 truncated isoform and p75, exhibited DNA binding properties similar to that of CR2CR3HD and CR3HD (89).

The binding of CDP1/Cux1 to an ATCGAT consensus binding site varies during the cell cycle. Binding was virtually undetectable in GO and early G1, but became very strong as cells reached S phase (27). Furthermore, CDP1/Cux1 DNA binding is inhibited by posttranslational modifications of either cut repeats or the cut homeodomain. Comparison with the sequence of *Drosophila cut* revealed that Cut repeats contain evolutionarily conserved consensus phosphorylation sites for protein kinase C (PKC) and casein kinase II (CKII). PKC (26) and CKII (28) were shown to phosphorylate Cut repeats, in vitro and in vivo. Phosphorylation of Cut repeats caused an inhibition of DNA binding and, consequently, the transcriptional repression activity of CDP1/Cux1. Furthermore, co-expression of CKII with CDP1 in mammalian cells or treatment of cells with an activator of PKC caused a decrease in DNA binding activity. Also, it has been reported that the CDC25A phosphatase is responsible for the increase in binding of CDP1/Cux1 in S phase due to dephosphorylation of CDP1/Cux1. Moreover, it was shown that the CDP1/Cux1 homeodomain is acetylated by the p300/CREB-binding protein-associated factor (PCAF) (73). Acetylation of CDP1/cux1 by PCAF is directed at conserved lysine residues near the homeodomain region and regulates CDP1/Cux1 function by inhibiting its DNA binding activity. Other results indicated that acetylation has no effect on DNA binding affinity and specificity of various isoforms of CDP1/Cux1, nor does it affect the phosphorylation by PKC (Muzzin-Unpublish data).

Another posttranslational modification of CDP1/Cux1, the phosphorylation of serine 1237 and 1270 by cyclin A-Cdk1, near and in the homeodomain, respectively will inhibit the DNA binding activity of the protein in the late S and G2 phases of the cell cycle (105). Furthermore, overexpression of cyclin A-Cdk1 will inhibit the DNA binding by a wild type protein but not by a mutant CDP1/Cux1 protein in which serines 1237 and 1270 were

replaced with alanine. These results suggest that CDP1/Cux1 is a physiological target of cyclin A-Cdk1 and that down-modulation of CDP1/Cux1 activity is important for cell cycle progression in late S and in G2.

The Cux1 knockout mice that lack the homeodomain and the C-terminal region, displayed a number of phenotypes in various organs including curly whiskers, growth retardation, delayed differentiation of lung epithelia, altered hair follicle morphogenesis, male infertility, and a deficit in T and B cells (37, 81, 111, 125). In contrast to the small size of the Cux1 knockout mice, transgenic mice expressing Cux1 under the control of the CMV enhancer/promoter displayed multi-organ hyperplasia and organomegaly. This raised the possibility that constitutive expression of Cux1 stimulated the proliferation of stem cells or the transient amplifying cells that are derived from them (69). Thus, from genetic studies both in Drosophila and the mouse, it is clear that the *CDP1/Cux1/Cut* gene plays an important role in the development and homeostasis of several tissues.

4. CDP2/Cux2, a first overview

A second *cut* homologue gene was recently discovered in all vertebrates and was named *CDP2* in human and *cux2* in mouse and chicken (63, 100, 119). The chromosomal loci that encode the already known CDP1/Cux1 and the new homologue CDP2/Cux2 have been designated Cut-like 1 (*CUTL1/Cutl1*) and Cut-like 2 (*CUTL2/Cutl2*) (100). In human, *CUTL1* is located on chromosome 7 and *CUTL2* on chromosome 12. In the mouse, *Cutl1* and *Cutl2* are both located on chromosome 5 but on different loci. This confirmed that *cux2* was not a splice variant of a single gene but the product of a single gene. The observation that *Cutl1* and *Cutl2* were genetically linked was unexpected since most previously identified members of families of diverged homeobox genes in mammals are dispersed, *e.g.* *En-1* and *En-2* on mouse chromosomes 1 and 5; *Evx-1* and *Evx-2* on mouse chromosomes 6 and 2; *Pax-3*, *Pax-4*, and *Pax-6* on mouse chromosomes 1, 6, and 2, respectively (100). *Cux1* and *cux2* may have arisen by *cis* duplication within a chromosome by unequal sister chromatid exchange as suggested for *Hox* genes (64). Thus, this demonstrates that *cux2* is a paralogous gene, which was derived from a gene duplication event. The identification of two paralogous gene within the same species verifies the existence of a family of *cut*-related homeobox genes in mammals (100).

The murine Cux2 protein contains 1426 amino acids with an estimated size of 155 kDa. Essentially, the nucleotide sequence of murine *cux2* is 65% identical to murine *cux1*. The amino acid sequence of Cux2 is 48% identical (63% similar) to murine Cux1 and 29% identical (50% similar) to *Drosophila cut*. Quite the reverse, the murine Cux1 amino acid sequence and its orthologues in human, rat and dog show an identity of 81-96% (86-97% similarity) to the *Drosophila* Cut. Cux2 contains, from the amino terminus to the C terminus a coiled-coil region, three Cut repeat sequences and a homeodomain all relatively spaced in the same way from each other inside the protein as in CDP1/Cux1 (refer to appendix, Fig. E). The Cux2 homeodomain is 75% identical (79% similar) to murine Cux1 and 47% identical (61% similar) to *Drosophila* Cut. However the Cux2 homeodomain shows less similarity to the archetypal Antennapedia sequence and Hoxb-7. The three 60-amino acids internal Cut repeats show some equivalence up to between 58-68% to *Drosophila* Cut and between 88-98% for Cux1 Cut repeats. Compared to Cux1, there is less similarity between each CR of Cux2 (60–67% identity). Also, the sequence similarity between Cux2, Cux1, and *Drosophila* Cut was highest near the carboxyl-terminal end of the Cut repeats (100). Cux2 and Cux1 share between the CR3 and the homeodomain a high cationic region that could represent a nuclear localization signal (43). Beside those entire domains, Cux2 shows three other high homology regions that have not been associated with a specific function, at in the amino terminus, one between CR1 and CR2 and one at the C-terminus. Cux2 does not have the high alanine and proline rich region at the C-terminus that was shown to be a repression domain in CDP1 and Cux1 (82). Murine Cux2 was generally not very similar to Cux1 beside all the domains described (33% amino acid identity).

Glutathione *S*-transferase (GST) fusion proteins containing the carboxyl-terminal Cut repeat and homeodomain of Cux2 could bind to an upstream regulatory element of the neural cell adhesion molecule *(Ncam)* promoter in a concentration-dependent manner. This promoter has previously been identified as a Cux1 target and the gene is also expressed in neural tissue (126). However, it remains to be verified whether it is a real target of Cux2. Moreover, the experiment was partially flawed by the fact that GST fusion proteins exists as dimers (50).

5. Cux2 expression in adult and developing mice

The expression pattern of *cux2* is remarkably different than that of *cux1*, which is expressed in most adult tissues with higher transcript levels in nephrogenic and urogenital system (3, 92, 100, 126, 131). In adult mice, *cux2* is expressed only in the nervous system (100), whereas it exhibits a dynamic expression pattern during mouse embryogenesis. *Cux2* expression becomes detected in a number of tissues during organogenesis including the olfactory epithelium, branchial arch and limb bud progress zones, roof plate, motor neurons, dorsal root ganglia and urogenital tissues (57). Interestingly, *cux2* expression in the maxillary and mandibular at E10.5 overlaps with *Msx2*, a downstream target of the bone morphogenetic proteins (BMP) (41, 49). This transcription factor is known to have a crucial role during lens induction and is expressed strongly in the optic vesicle, weakly in the surrounding mesenchyme and the surface ectoderm. *Cux2* expression also overlaps with both *Bmp4* and *Msx2* in the progress zone of the limb bud. Altogether, it seems that Cux2 may be regulated by BMP signaling (57). Expression of *cux2* in the late embryogenesis, E13.5 to E16.5, was restricted to neural tissue in the developing central and peripheral nervous systems (100). The highest level of expression of *cux2* was in the telencephalon and trigeminal ganglion. Lower levels of expression were detected in spinal cord. *Cux2* transcripts were also detected in peripheral ganglia of the glossopharyngeal (*IX*) nerve (100). The highest expression of *cux2* was in the surface of the cortical plate, a cell layer of the developing telencephalon that contains differentiating, postmitotic neurons. Much lower levels of *cux2* were detected in the ventricular and subventricular zones (*vz*), which contain proliferating neuroblasts (100). The earliest localization of *cux2* was in the developing motor neuron in the ventral horn of the neural tube, and in the roof plate (rp) of the dorsal neural tube (57). *Cux2* was expressed in other specific neuron cell types such as V3 interneurone and commissural neuron. In adult mice the highest expression of *cux2* was in the thalamus and the limbic system. In the thalamus *cux2* was highly expressed in specific nuclei including the laterodorsal (*ldvl*), ventroposterior (*vp*), lateroposterior (*lpmr*), mediodorsal (*mdl*), and centromedian (*cm*) nuclei. (refer to appendix, Fig. D) Within the limbic system, *cux2* is highly expressed in the hippocampus proper (*h*), dentate gyrus (*dg*), and amygdala (*am*). Lower levels of *cux2* expression were detected in the cerebral cortex (*co*) and external capsule (*ec*) (100).

6. Cux2 in chick limb development

Drosophila and vertebrates have been found to be quite similar throughout the signaling process in limb development. In the case of Drosophila *cut*, it was shown that it plays a role in the dorsal-ventral (D-V) boundary cells during the development of the wing imaginal discs (56). In more details, it is thought that Cut is one of the downstream effectors of Notch, and would act during late larval and pupal stages to maintain Wingless expression and suppress that of Serrate and Delta in the presumptive wing margin. In vertebrates this function seams to be fulfilled by two genes, *cux1* and *cux2*. During chick limb outgrowth *cux1* and *cux2* are expressed in different compartments of the limb bud. *Cux1* is present in both the dorsal and ventral limb ectodermal cells bordering the Apical Ectodermal Ridge (AER) (119). *Cux2* is expressed in the pre-limb lateral plate mesoderm, posterior limb bud and flank mesenchyme, a pattern reminiscent of the distribution of polarizing activity (119). *Cux1* expression is regulated by AER signals and it will control the AER formation. *Cux1* is maybe involved in the establishment of the AER at the distal tip of the limb ectoderm (119). On the other hand, the pattern of expression of *cux2* indicates that it could be involved in the establishment of the Zone of Polarizing Activity (ZPA) (119). It is a group of cells located at the posterior region of the developing limb bud that control the posterior patterning of the vertebrate limb (106). It was demonstrated that *cux2* can be regulated by both retinoic acid (molecule involved in polarizing activity) and a molecular component of the ZPA, the *sonic hedgehog* gene (119). These results suggest that *cux2* may be a nuclear target of some of the known pathways implicated in anteroposterior patterning of the vertebrate limb (119). Overall, chick *cux1* and *cux2* may act by modulating proliferation versus differentiation in the limb ectoderm and the region of polarizing activity (ZPA), respectively (119). Overall this suggests that the role of *Drosophila cut* in the development of the limb has been conserved in evolution by these two *cut* homologue, *cux1* and *cux2*. Furthermore, an example of a possible dynamic evolution of the Cut gene in vertebrate can be observed by the fact that in the chick embryo *cux1* (and not *cux2*) expression in the craniofacial primordial is analogous to that of *cux2* in the mouse. Elsewhere, we can observe similar expression of *cux2* in both the chick and the mouse in the developing whisker follicles, the face and the limbs (57).

7. CDP2 and susceptibility to bipolar disorder

It was reported that the bipolar affective disorder, a genetically complex psychiatric disease, and the Darier disease, a dominant skin disorder with a neuropsychiatric component, were cosegregate. Furthermore, the human *CUTL2* is mapped to the same chromosomal region than those two diseases, at 12q23-q24.1. An increased level of the NCAM protein is found in the bipolar affective disorder (98). The neural cell adhesion molecule (NCAM) is involved in many processes of hippocampal development, including axon fasciculation, growth cone guidance, target recognition, and, in the adult brain, ongoing synaptic plasticity (132). Furthermore, as a glutathione-*S*-transferase (GST) fusion proteins containing the carboxyl-terminal Cut repeat and homeodomain of the mouse Cux2 could bind to an upstream regulatory element of the Ncam promoter sequence in a concentration-dependent manner it was suspected that mutations in the *CDP2* coding sequence could be implicated in the bipolar disorder. However, mutational analysis of the *CDP2* gene revealed no evidence for the involvement of variants within the *CUTL2* coding region that could produce susceptibility to bipolar disorder (63).

8. Goal of this study

A new Drosophila *cut* homologue has been recently identified in vertebrates, *cux2*. Both genes exhibit a dynamic expression pattern during mouse embryogenesis. In the adult mice, however, *cux1* is expressed in a wide range of tissues whereas *cux2* is only expressed in the nervous system and more particularly in certain neural tissues that are involved in processing sensory information. Genetic studies in *Drosophila melanogaster* indicated that *cut* plays an important role in determining cell-type specificity and that was most thoroughly demonstrated in peripheral nervous system. This particular function of *cut*, which is to specify neural identity, may be phylogenetically conserved in mammals and fulfilled by *cux2*. Furthermore, while the *cux1* knockout exhibited defects in several tissues, it was able to complete embryonic development and, in the appropriate genetic background, to survive. This result was surprising as deletion of the *cut* gene in Drosophila caused embryonic lethality. One possible explanation is that the *cux1* and *cux2* genes may be partially redundant in higher vertebrates. Another possibility is that in the context of the *cux1* knockout, *cux2* became more widely expressed and was able to partially complement the *cux1* defects. Apart from the detection of the transcript in mice, little is known about

the Cux2 protein. My goal was to characterize the biochemical properties of the Cux2 protein and to generate reagents that would enable the study of Cux2 protein expression in vivo and in tissue culture. The biochemical properties that were to be investigated included the DNA binding activities of the protein as a whole, and of various combination of its DNA binding domains. Other properties include the intracellular localization of the Cux2 protein and its ability to be proteolytically processed.

Materials and Methods

Cell culture and transfections

UW1, UW2, UW3 and DAOY are human medulloblastoma cell lines, the exact cells of origin in not known (122). SK-N-SH, SK-N-AS, SK-N-F1, SH-SY5Y, NGP and N1E-115 and B104 are neuroblastoma cell lines (10, 17, 101, 103, 117, 118). Neuroblastoma comes from the adrenal medulla or other sites of sympathetic nervous tissues. They are all human cell lines beside B104 which is a rat and N1E-115 which is a mouse cell line. NG-108b is a mouse/rat hybrid between a neuroblastoma and glioma cells (48). OLN-93, CG4, C6, U251-MG, U343 MG-A, U87-MG are glioblastoma and glial cells (8, 22, 121, 124, 134). OLN-93 is an oligodendrocyte glial cell type and the CG4 is a primary oligodendrocyte cell line. C6, U251-MG, U343 MG-A and U87-MG come from astrocyte tumor glial cell types. The BC3H1 cell line was established from a mouse brain tumor induced by nitrosoethylurea. They retain properties characteristic of muscle cells (107). FR3T3 and NIH3T3 are respectively rat and mouse fibroblastic cell lines. The HS 578T cell strain was derived from a carcinoma of the breast; the cells come from the mammary gland. FR3T3, HS 578T, NIH3T3, OLN-93, B104, C6, UW1, UW2, UW3, DAOY, SK-N-SH, SK-N-AS, SK-N-F1, U251-MG, U343 MG-A, U87-MG and NGP were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) containing 1% (v/v) glutamine and 1% (v/v) penicillin/streptomycin. All the cell lines were grown at 37°C in a 5% CO2 environment, except the OLN-93 and B104 that needed a 10% CO2 environment. The CG4 cell line is maintained undifferentiated in N1 medium supplemented with Biotin and the B104 supplemented N1 in a 70:30 ratio. N1 medium contains complete DMEM (with non-essential amino acids, pyruvate), N1 supplement (Sigma), 1% (v/v) penicillin/streptomycin and biotin (stock 1ug/uL) 5 uL in 500mL. The B104 supplemented N1 medium comes from a conditioned B104 culture with N1 medium for 72h. The CG4 were grown at 37°C in a 10% CO2 environment. The SH-SY5Y, NG-108b and BC3H1 cell lines were grown in accordance to the ATCC's instructions (CRL-2266; HB-12317; CRL-1443).

To differentiate some of the cell lines, serum starvation and retinoic acid (RA) treatment was used. The N1E-115 was grown to a 75% confluence followed by a serum

starvation for 3 days. The SK-N-AS, SK-N-F1, SH-SY5Y and the NGP cell lines were differentiated with RA (38, 75, 42). The SK-N-AS cells were split to obtain a 60% confluence. The next day, 50 μM of RA was added to the media for 3 days. The SK-N-F1 cells were split to obtain 60% confluence. The next day, 100 μM of RA was added to the media for 5 days. The media was changed with the addition of the same concentration of RA after 3 days. The SH-SY5Y and the NGP cells were split to obtain 60% confluence. The next day, 50 μM of RA was added to the media for 4 days.

The transient transfections were performed with GeneJuice (Novagen) in accordance with the manufacturer's instructions. Briefly, HS 578Tand NIH3T3 were plated to obtain 80% confluence. The next day, the cells were washed with PBS and 10 mL of medium without serum was added. 10 μg of plasmid DNA (in a total of 50 μL of PBS) was mixed to a solution containing 20 μL of GeneJuice reagent with 930 μL of media for 30 min. Then, this mix was added to the cells. After 2h, 10 mL of 20% FBS complete media was added.

Preparation of cellular extracts

Nuclear extracts were prepared according to the procedure of Lee et al., excepting that the nuclei were obtained by submitting cells to 3 freeze-thaw cycles in buffer A (10 mM HEPES [pH 7.9], 10 mM KCl, 1.5 mM MgCl2, 1 mM dithiothreitol [DTT] (Lee-1988). The nuclear pellet was then resuspended again in buffer C (20 mM HEPES [pH 7.9], 25% glycerol, 1.5 mM MgCl2, 420 mM NaCl, 0.2 mM EDTA) and incubated at 4°C for 30 min. After 15 min of centrifugation, the supernatant was collected. Buffers A and C were supplemented with protease inhibitor mix tablets purchased from Roche. For the nuclear pellet extract, it was resuspended again in 100 μL of 1X of loading buffer (5X: 2% SDS, 10% glycerol, 50 mM Tris-HCl, pH 6.8, 0.1% bromophenol blue). Total extracts were prepared by applying buffer X (50 mM HEPES [pH 7.9], 0.4 M KCl, 4 mM NaF, 4 mM Na3VO4, 0.2 mM EGTA, 0.2 mM EDTA, 0.1% NP-40, 10% glycerol, 0.5 mM DTT, protease inhibitor mix tablet (Roche)) to a monolayer plate. After 10 min of incubation on ice, the resulting slurry was centrifuged for 15 min at 4°C and the supernatant was collected.

Plasmids description

The plasmids for expression of histidine-tagged Cux2 fusion proteins were prepared by inserting various fragments from the mouse Cux2 cDNA (GenBank, accession No. U45665) into the bacterial expression vectors pET-15b (Novagen). CR23HD: nt 2763 - 3949 into the BsrG1 site of pET-15b. CR3HD: nt 3191 - 3949 into the Xmn1 site of pET-15b. The CR1CR2 fusion protein was amplified by PCR to insert an Xho1 site at the 5'end and a BamH1 site at the 3'end: nt 1745 - 2980. This product was digested and inserted in the Xho1 and BamH1 site of pET-15b.

The expressing vectors for GST-Cux2 fusion protein used to raise antibodies were prepared by inserting a choice of different antigenic regions restricted to Cux2 in PET15-b plasmid. These regions were produced by PCR amplification (See Oligonucleotides). The PCR product from aa 48 to 180, BamHI-EcoRI was introduced into XbaI of pET-15b; aa 356 to 415, BamHI-EcoRI into Styl site of pET-15b; aa 653 to 680, BamHI-EcoRI into Styl site of pET-15b; aa 748 to 771, BglII-EcoRI into BglI site of pET-15b; aa 902 to 945, BamHI-EcoRI into NheI site of pET-15b; aa 1177 to 1401, BamHI-EcoRI into Styl site of pET-15b.

The original Cux2 plasmid was provided in the PKS plasmid. To transfer Cux2 into the pMX139 vector, the Cux2 insert was digested with Not1 and Kpn1 in PKS/Cux2 and the pMX139 vector was digested with Sma1. Those products were ligased. To transfer Cux2 coding sequence in pMX139/Myc-HA, this vector was digested with Mlu1 and Sma1. For the insert, a PCR was done on the pMX139/Cux2 to insert an Mlu1 site on the 5'end, to be in frame with the KOSAK and myc tag. Also, the stop codon of Cux2 was removed to be in frame at the end with the HA tag. The PCR product was obtained and digested with Mlu1 and Sma1. The portion of the PCR product was replaced with the portion of the original DNA (to avoid the sequencing of the whole Cux2 sequence. The vector was obtained by digesting pMX/Myc/Cux2/HA with Bcl1 and EcoRV and the insert was obtained by digesting pMX/Cux2 with the same enzymes. The CDP2 expressing vector used is in pMX139/Myc/CDP2/HA. The CDP1 expressing vector used was pMX139/Myc/CDP1/HA as previously described (89). The Cux1 expressing vector is in pMX139.

Oligonucleotides

The sequences of oligonucleotides used in this study for EMSA experiment are as follows:

29FP ATCG: "CGATATCGAT", TCGAGACGAT**ATCGAT**AAGCTTCTTTTC.

29FP: "ATCGAT", TCGAGACGGT**ATCGAT**AAGCTTCTTTTC.

CDA: "CCAAT–CGAT," TCGAGAAAAGAACAA**CCAATCACCGAT**C.

For the production of the antibodies, each region was amplified by PCR. Here are the nucleotide regions used as primers for each peptide amplification: 48: 416 to 441+ BamH1 and 698 to 577+EcoR1. 356: 1340 to 1364+BamH1 and 1516 to 1494+EcoR1. 653: 2516 to 2542+BglII and 2587 to 2565+ EcoR1. 748: 2978 to 2997+ BamH1 and 3109 to 3085+ EcoR1. 902: 3803 to 3831+ BglII and 4477 to 4453+ EcoR1 1177: 2231 to 2253+BamH1 and 2314 to 2290+ EcoR1.

Here are the primers for the insertion of Cux2 in pMX/MH:

5'-ACTGACGCGTGTAGCTAAGGTGCTGAAGAGCTTCCAGGC-3';

5'-AGCTGATATCGAATTCCCACTCCAGGACCTCTTCCCG-3'

For the production of Cux2 CR1CR2 Histidine-tagged fusion protein a PCR product was used, here are the primers:

3'-GACCTGTGCCGACTCTAGCG-5' and 3'-GATGCAGTTGTGGCTGTCGGAC-5'.

For the CDP2 cDNA specific amplification, the oligos used are:

5'-CATCACTGGCCACCGCGAG-3'and

5'-TGAGTACATCTTCCCGCTCAACGA3'

For the mouse Cux2 cDNA specific amplification, the oligos used are:

5'-CCACTCGGGTCAAAGTCAGGATCA-3'and

5'GTGAGCATCCCCAACGGCACA-3'

For the rat Cux2 cDNA specific amplification, the oligos were prepared upon the EST sequence: EST450556 (GenBank Accont: BF285965), they are:

5'- TTTTTGCTCCTGCCTTCCCG -3' and

5'- AATTCTGAAGCTGGGGCAAGGGAC -3'

For the MEF genotyping, the following oligos were used:

(1) 5 –TGCTCATCCACCTGCCTCAATGTC-3 , (2) 5 –ATCCATCAGCT

TCTCCACATTGTT-3 , and (3) 5 –TCCTGTAGCCAGCTTTCATCAACA-3 .

Expression and purification of Histidine-tagged (His-tag)-Cux2 fusion proteins

The different his-tag Cux2 constructs were introduced into the BL21 (DE3) strain of *Escherichia coli* and induced with isopropyl-b-D-thiogalactopyranoside (IPTG). The fusion proteins were purified by affinity chromatography following procedures recommended by the suppliers (Invitrogen).

Expression and purification of glutathione-S-transferase (GST)-Cux2 proteins for antibody production

The vectors encoding the indicated region of GST-Cux2 fusion proteins were introduced in the DH5 strain of *E. coli* and protein expression was induced with 1 mM IPTG for 1.5 hours. The fusion proteins were purified on glutathione sepharose (Amersham Pharmacia Biotech) according to the manufacturer's instructions.

Electrophoresis Mobility Shift Assay (EMSA)

EMSA were performed with either 10 ng of purified fusion proteins or 5 μg of nuclear extract from mammalian cells. The samples were incubated at room temperature for 5 minutes, in a final volume of 30 μl of 25 mM NaCl, 10 mM Tris, pH 7.5, 1 mM MgCl2, 5 mM EDTA, pH 8.0, 5 % of glycerol, 1 mM of DTT, with 30 ng of poly dIdC and 3 μg of BSA as nonspecific competitors. End-labeled (^{32}P) double stranded oligonucleotides (~10 pg) were added and further incubated for 15 minutes at room temperature. Samples were loaded on a 4% polyacrylamide gel (30:1) and separated by electrophoresis at 8V/cm in 0.5X TBE or 1X Tris-glycine running buffer. The gels were dried and visualized by autoradiography.

RNA isolation and RT-PCR assays

Total RNA was isolated from different mouse tissues, rat brain (from Dr McPherson), human cortex brain (from Dr Hamel) and mouse embryo fibroblast (MEF) cells using Trizol regent according to the manufacturer's instructions (Invitrogen). RT-PCR was performed with 20 μg of total RNA with Superscript II reverse transcriptase

(InVitrogen) using oligo dT as primers. Boiling for 10 minutes stopped the reaction. An aliquot of 3 µl was used as a template for polymerase chain reaction (PCR) using either Cux2, CDP2, CDP1 and GAPDH specific primers with the Taq DNA polymerase (Invitrogen). For the mouse CDP2 cDNA specific amplification the PCR was performed in a final volume of 30 µl, containing approximately 1 ng cDNA, 7,5 % DMSO, 3 µl 10x PCR buffer (-)MgCl2 (Invitrogen), 0,9 ρM of each primer, 0,6 mM dNTP, 1,5 mM MgCl2, 0,25 µl of Taq DNA polymerase (Invitrogen). An initial step of 3 min at 95°C was followed by 40 cycles of 45s denaturation at 94°C, 45s of annealing at 61°C, 1 min of extension at 72°C, followed by a final extension step of 7 min at 72°C. For the mouse Cux2 cDNA specific amplification the PCR was performed in a final volume of 30 µl, containing approximately 1 ng cDNA, 3 µl 10x PCR buffer (-)MgCl2 (Invitrogen), 0,9 ρM of each primer, 0,6 mM dNTP, 1,25 mM MgCl2, 0,25 µl of 0,25 µl of Taq DNA polymerase (Invitrogen). An initial step of 3 min at 95°C was followed by 30 cycles of 45s denaturation at 94°C, 45s of annealing at 61°C, 2 min of extension at 72°C, followed by a final extension step of 7 min at 72°C. 5 µL of the reaction was loaded on either 0,8% agarose gel or 4% acrylamide DNA gel.

Cux2 antibodies and Western blot analysis

To generate polyclonal antibodies against various regions of Cux2, fusion proteins were prepared in which the glutathione-S-transferase (GST) protein was fused to various regions of Cux2: antibody #48, residues 48-108; #356, 356-415; #653, 653-680; #748, 748-771; #902, 902-945. Two rabbits were injected with 500 µg of purified bacterial fusion proteins. The animals were boosted twice with 250 µg of proteins, and serum was collected 10 days after the last boost. For the antibodies purification, see the antibodies purification section. For Western blot analysis, protein extracts were recovered as described above. The loading buffer was added to the extract to reach a 1X concentration (5X: 2% SDS, 10% glycerol, 50 mM Tris-HCl, pH 6.8, 0.1% bromophenol blue) and separated by electrophoresis on 6% polyacrylamide gels. Western blot analysis with the 861 Cux1 antibody was performed as previously described (Coqueret-1998). Western blot analysis for Myc and HA epitopes: the membranes were blocked with 10% milk. For the Myc blot, 2% bovine serum albumine (BSA) in TBST (10 mM Tris pH8, 150 mM NaCl , 0.1% Tween) was added for 3h at

room temperature. Myc or HA antibodies were incubated with the membrane in TBST for 1 hour at room temperature. After 4 X 10 min washes with TBST, the mouse secondary antibodies (Santa Cruz biotechnology) was added to the membranes in TBST and incubated for 45 min at room temperature. Following 4 X 10 min washes with TBST, proteins were visualized with the ECL system from Amersham. For the Cux2 antibodies, the PVDF membranes were blocked with 10% milk and 2% BSA for 3h followed by incubation for 1h at room temperature with the antibody diluted 1/2000 in TBST. After 4 X 10 min washes with TBST, the rabbit secondary antibodies (Santa Cruz biotechnology) were added to the membrane in TBST and incubated for 45 min at room temperature. Following 4 X 10 min washes with TBST, proteins were visualized with the ECL system from Amersham.

Antibodies purification

5 to 10 μg of the purified GST-Cux2 peptides and the GST protein alone, were dialyzed against the coupling buffer (0.1 M NaHCO3 + 0.5 M NaCl pH 8.3) overnight at 4°C. 0,5 g of CNBr-activated Sepharose ^{TM} 4B were activated following procedures recommended by the suppliers (Amersham Pharmacia Biotech AB). The beads and the antigen solution were mixed and rotated at 4°C overnight. The beads were then washed with 50 mL of coupling buffer. The remaining active groups were blocked by transferring the beads to 0,1 M Tris HCl (pH 8,0) for 2h. The columns were washed with 15 mL of 10 mM Tris (pH 8) containing 0,5 M NaCl, followed by 15 mL of 0,1 M Glycine (pH 2,5) containing 0,5 M NaCl, followed by 15 mL of 10 mM Tris (pH 8) containing 0,5 M NaCl, followed by 15 mL of 100 mM Triethylamine (pH 11.5) and then by 15 mL of 10 mM Tris (pH 8) containing 0,5 M NaCl. 4 to 6 mL of the antiserum was diluted to a final volume of 10 mL in PBS. The serum was passed three times through the GST column. The GST column was then eluted with 10 mL of 0,1M Glycine (pH 2,5) and washed twice with 10 mL of 10mM Tris (pH 7,5) containing 150 mM NaCl and 0,02 % sodium azide and stored. The flow through was applied to the antigen column three times. The columns were washed with 50 mL of 10 mM Tris (pH 7,5) containing 0,5 M NaCl. The antibodies were eluted with 6 mL of 0,1 M Glycine (pH 2,5) in six 1 mL fractions containing 100 μL of 1 M Tris-HCl (pH 8). The columns

were washed with 15 mL of 10 mM Tris (pH 8) containing 0,5 M NaCl. Another elution was done with 100 mM Triethylamine with 100 µL of Tris-HCl (pH 7.5) in the elution tubes. The columns were washed as the GST column and stored. The antibodies were applied to the Centricon YM-30 (Millipore), as the supplier recommendation. The antibodies were diluted to 0,5 mg/mL by adding 1/10 of 10% BSA and sodium azide 0,02%. The antibodies were stored at -80°C.

Cellular localization of Cux2

NIH3T3 cells on glass coverslips were transfected with a vector expressing Myc/Cux2/HA and Myc/CDP2/HA. The next day, the coverslips were transferred to a 24-wells plate previously filled with PBS. All the following steps were performed at room temperature (RT). PBS was removed and the coverslips were fixed with 3.7% freshly prepared paraformaldehyde for 10 min. The coverslips were then rinsed twice with PBS. The quenching solution (50 mM of NH4Cl) was added for 10 min followed by two PBS washes. The solubilizing mix (0.5% Triton-X, 5% FBS in PBS) was added for 10 min followed by two PBS washes. The coverslips were next incubated for 10 minutes in the solubilizing mix. The primary antibody was diluted in the solubilizing mix to the desired concentration and incubated on the coverslips for 1hr at RT. The primary antibody was either against the HA tag or Myc tag. The primary antibody incubation was followed by five PBS washes. The secondary antibody (Alexa Fluor α 488 (green)) goat anti-mouse from Molecular Probes was then added at a concentration of 1:1000, in the dark, for 30 min. Five PBS washes followed. Finally, the coverslips were mounted on slides and cells were visualized with incident light fluorescence with light microscopy.

Cux1$^{-/-}$ mouse embryo fibroblast (MEF) cells production

The mouse embryos used in this study were obtained from the Cux$^{-/-}$ mice stain (37). The mouse embryo fibroblasts were prepared according to the volume: Manipulating the Mouse Embryo, a laboratory manual. Hogan B, Beddington R, Costantini F and Lacy L. Second Edition, Cold Spring Harbor Laboratory Press, 1994. The mRNA was prepared and an RT-PCR assays were done, as describe above. The

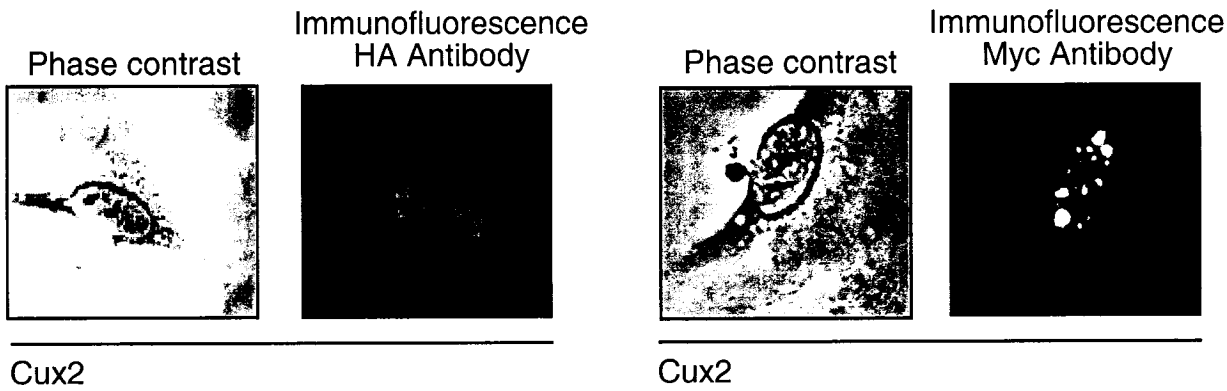
genotyping was achieved by PCR analysis (See oligonucleotide for sequences). The wild type and *Cutl1*$^{-/-}$ alleles gave rise to PCR products of 615 bp (primers 1/2) and 962 bp (primers 1/3), respectively.

Figure 1. CDP2/Cux2 are localized in the nucleus.

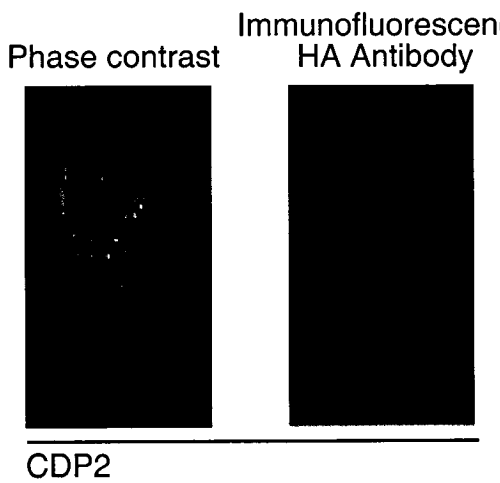
(A,B) NIH 3T3 cells on glass coverslips were transfected with a vector expressing either Myc/Cux2/HA or Myc/CDP2/HA. One day later, indirect immunofluorescence was performed with the HA or Myc antibodies (Ab) as described in materials and methods. The coverslips were mounted on slides and cells were visualized with incident light fluorescence with light microscopy. The left panel shows the phase contrast image and the right panel shows the immunofluorescence with either the HA or Myc Ab.

(C) Schematic representation of Myc/Cux2/HA and Myc/CDP2/HA construct used in the transient transfection are presented. The regions of CDP2/Cux2 recognized by the HA and Myc Ab are shown.

A


Phase contrast
Immunofluorescence
HA Antibody
Cux2
Phase contrast
Immunofluorescence
Myc Antibody
Cux2

B


Phase contrast
Immunofluorescence
HA Antibody
CDP2

C

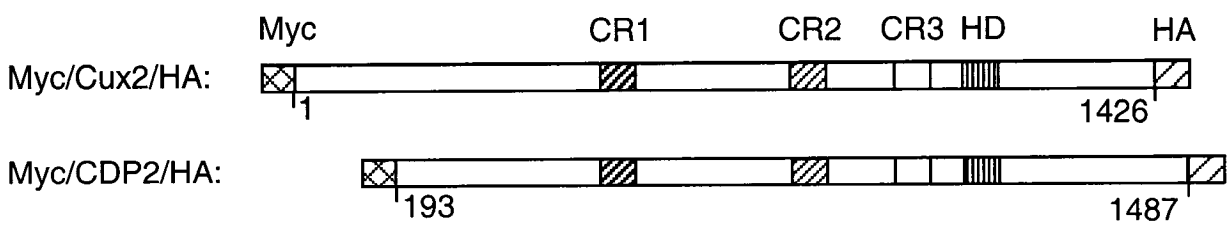
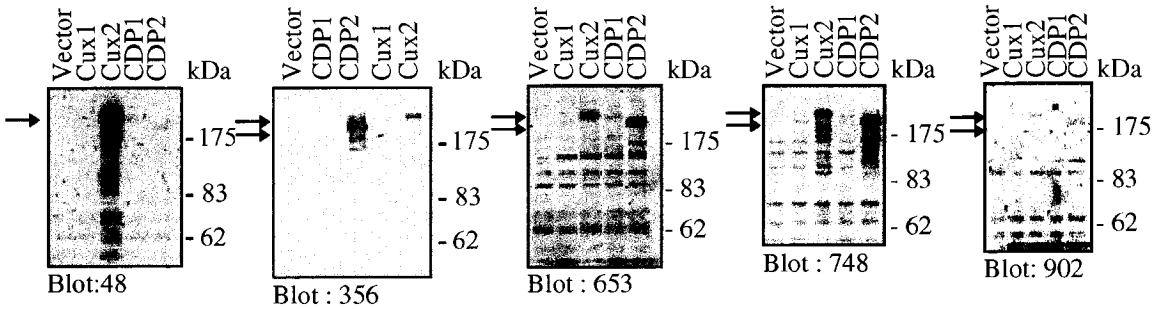

Myc
CR1
CR2
CR3
HD
HA
Myc/Cux2/HA:
1
1426
Myc/CDP2/HA:
193
1487

Figure 2. Five novel antibodies recognize Cux2 in Western blot.

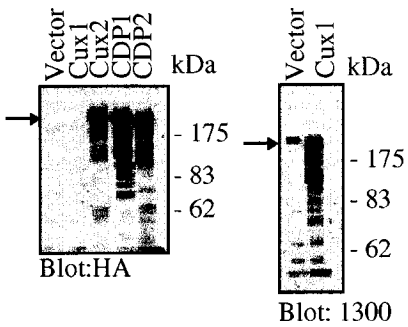
(A) Results of immunoblotting with Cux2 antibodies. HS 578T were transfected with 10 μg of vector DNA expressing either CDP2, Cux2, CDP1, Cux1 or the parental vector. 5 μg of proteins from nuclear cell extracts were resolved on a 6% SDS polyacrylamide gel and transferred to a PVDF membrane followed by immunoblotting using either 48, 356, 653, 748 and 902 Cux2 antibodies (Ab). The control blots were done with the HA Ab and the1300 Cux1 Ab. The arrows show the Cux1, CDP1, CDP2 and Cux2 proteins.

(B) Map of Cux2, indicated below the regions used to raise the antibodies. The expression vectors are shown.

A


Vector
Cux1
Cux2
CDP1
CDP2
kDa
175
83
62
Blot:48
Blot : 356
Blot : 653
Blot : 748
Blot: 902

Controls:


Vector
Cux1
Cux2
CDP1
CDP2
kDa
175
83
62
Blot:HA
Blot: 1300

B

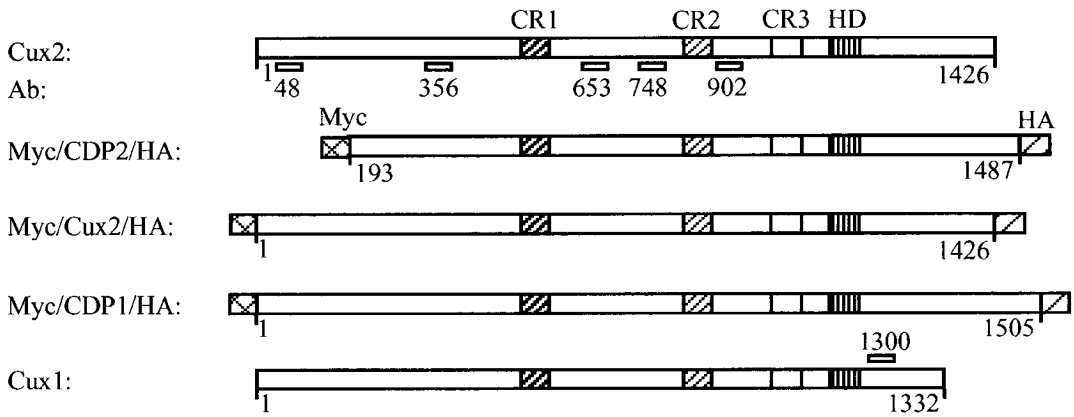

Cux2:
Ab:
CR1
CR2
CR3
HD
1
48
356
653
748
902
1426
Myc/CDP2/HA:
Myc
HA
193
1487
Myc/Cux2/HA:
1
1426
Myc/CDP1/HA:
1
1505
Cux1:
1300
1
1332

Figure 3. Five novel antibodies recognize Cux2 in electrophoretic mobility shift assays (EMSA).

(A-B) EMSA, NIH 3T3 cells were transfected with 10 μg a vector expressing the Myc/Cux2/HA, Myc/CDP2/HA or the Cux1 protein. The nuclear extracts were prepared and 1 μg was incubated with oligonucleotides containing a universal CDP1/Cut consensus-binding site 29FP ATCG. The supershifts were done with the five Cux2 Ab. The controls were done with the 861 Cux1 Ab. The arrows show either the protein-DNA complexes or the supershift of the appropriate antibody. Untrans. : Untransfected NIH3T3 cells extract.

(C) Map of Cux2, indicated below the regions used to raise the antibodies. The expression vectors are shown.

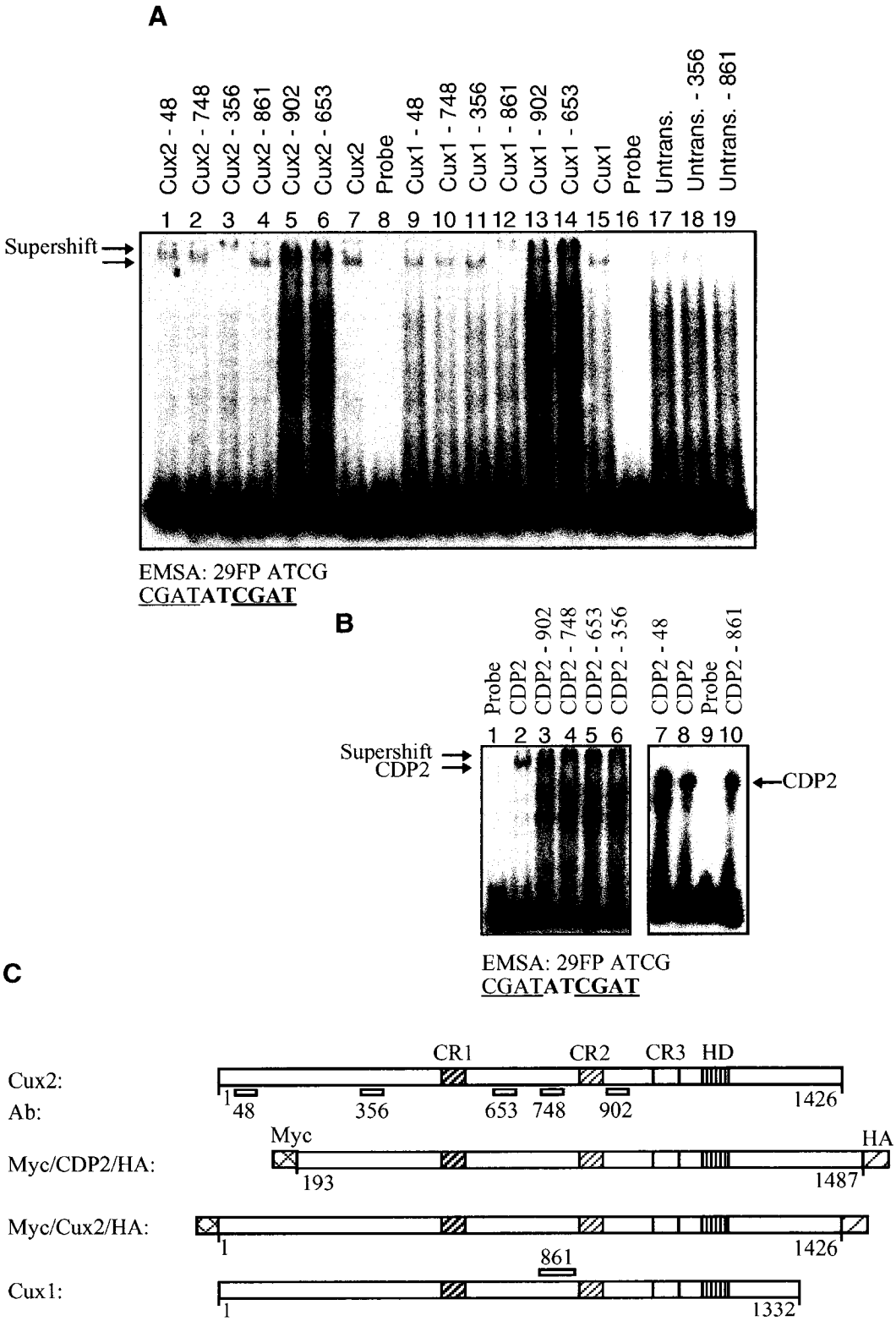

A
Cux2 - 48
Cux2 - 748
Cux2 - 356
Cux2 - 861
Cux2 - 902
Cux2 - 653
Cux2
Probe
Cux1 - 48
Cux1 - 748
Cux1 - 356
Cux1 - 861
Cux1 - 902
Cux1 - 653
Cux1
Probe
Untrans.
Untrans. - 356
Untrans. - 861
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Supershift
EMSA: 29FP ATCG
CGATATCGAT
B
Probe
CDP2
CDP2 - 902
CDP2 - 748
CDP2 - 653
CDP2 - 356
CDP2 - 48
CDP2
Probe
CDP2 - 861
1 2 3 4 5 6 7 8 9 10
Supershift
CDP2
CDP2
EMSA: 29FP ATCG
CGATATCGAT
C
CR1
CR2
CR3
HD
Cux2:
Ab:
1
48
356
653
748
902
1426
Myc
HA
Myc/CDP2/HA:
193
1487
Myc/Cux2/HA:
1
1426
861
Cux1:
1
1332

Figure 4. Cux2 DNA binding domains function in pairs and bind to specific sequences like Cux1.

(A) Cux2 histidine-tagged (His-tag) fusion proteins are schematically represented. The fusion proteins were expressed in bacteria as His-tag fusion proteins. Radiolabeled oligonucleotides encoding the ATCGAT site with various flanking sequences were incubated with decreasing amounts of purified fusion proteins at room temperature for 15 minutes (min) and resolved on a nondenaturing polyacrylamide gel that was dried and autoradiographed.

On the left (B,E and H), the original 29FP oligonucleotides containing the ATCGAT site also includes the sequence CGGT, which closely resembles the CGAT consensus half-site for CR1CR2. In the middle (C, F, and I), point mutations were introduced in the second half-site to produce a perfect CGAT half-site. On the right (D, G, and J) the NF-Y consensus binding site (CCAAT) flanked with a CGAT site was used; since this sequence does not contain ATCGAT, it is not expected to be recognized as well by proteins that do not contain the homeodomain.

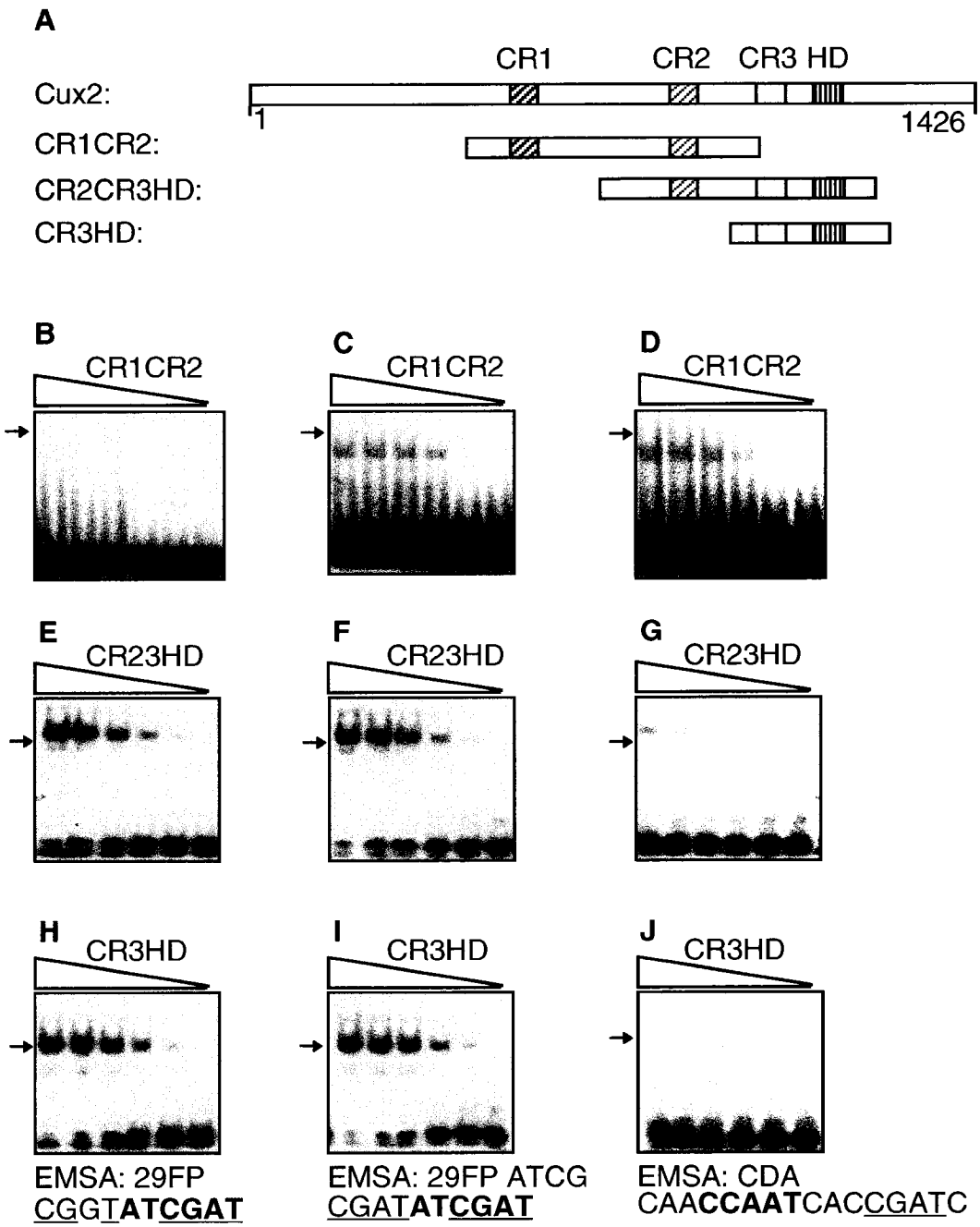

A
CR1
CR2
CR3 HD
Cux2:
1
1426
CR1CR2:
CR2CR3HD:
CR3HD:
B
CR1CR2
C
CR1CR2
D
CR1CR2
E
CR23HD
F
CR23HD
G
CR23HD
H
CR3HD
I
CR3HD
J
CR3HD
EMSA: 29FP
CGGTATCGAT
EMSA: 29FP ATCG
CGATATCGAT
EMSA: CDA
CAACCAATCACCGATC

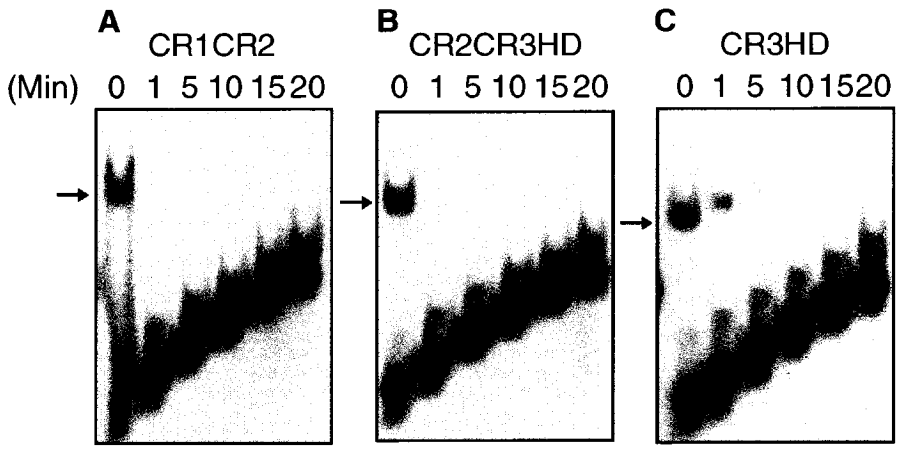
Figure 5. Off and on rates of Cux2 CR1+2, CR23HD and CR3HD.

(A-C) 100 ng of the indicated fusion protein were used in panel A and B, while 25 ng was used in panel C. Proteins were incubated with radiolabeled oligonucleotides at room temperature until the equilibrium was reached (15 min). 1000 fold molar excess of unlabeled oligonucleotides was added, and at the indicated time points, aliquots of the mixture were taken and analyzed in EMSA.

(D-F) 100 ng of the indicated fusion protein were used in panel D and E, while 25 ng was used in panel F. Proteins were mixed with radiolabeled oligonucleotides. The incubation took place at 4°C and at the indicated time points, aliquots of the mixture were taken and resolved on a nondenaturing polyacrylamide gel. For all the panels, the radiolabeled 29FP ATCG oligonucleotides containing the ATCGAT and CGAT sequences was used.

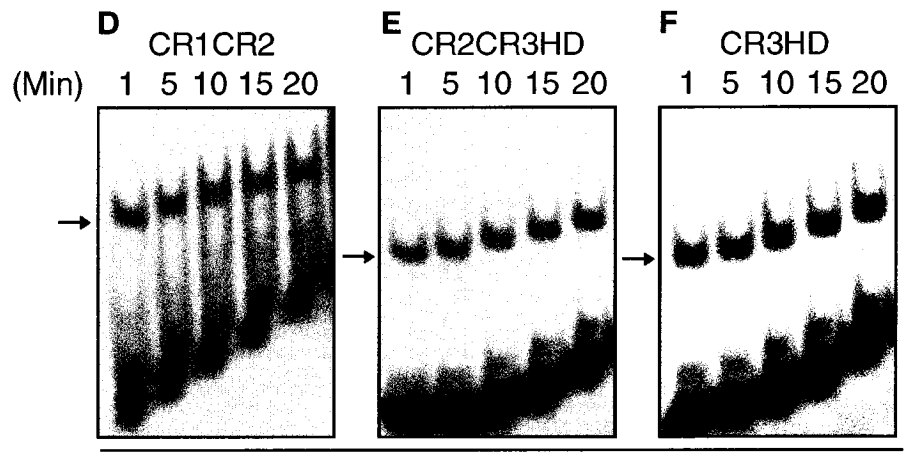
(G) Cux2 his-tag fusion proteins are schematically represented.

OFF RATES:



EMSA: 29FP ATCG
CGAT**ATCGAT**

ON RATES:



EMSA: 29FP ATCG
CGAT**ATCGAT**

G

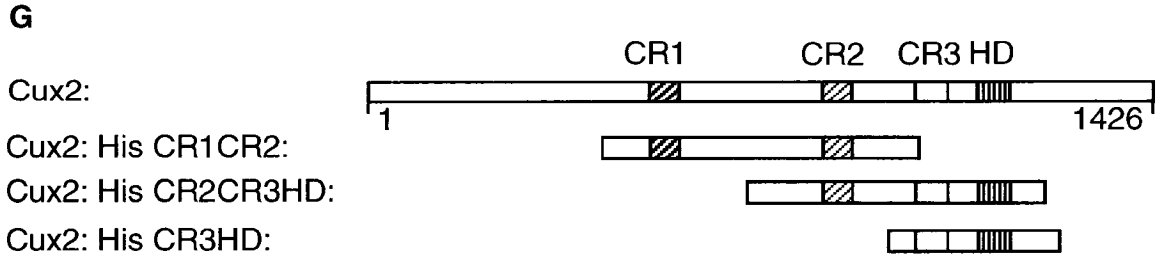


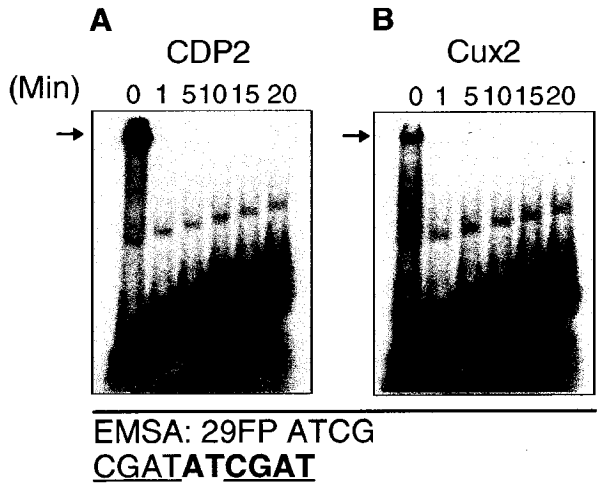
Figure 6. Off and on rates of full length CDP2 and Cux2.

(A,B) Off rates of full length CDP2/Cux2. NIH 3T3 were transfected with 10 μg of either the full length CDP2 or Cux2 construct. 20 ng of the nuclear extract was incubated with radiolabeled 29FP oligonucleotides at room temperature until the equilibrium was reached (15 min). 1000 fold molar excess of unlabeled oligonucleotides was added, and at the indicated time points, aliquots of the mixture were taken and analyzed in EMSA.

(C,D) On rates of full length CDP2/Cux2. 20 ng of either Cux2/CDP2 nuclear extract was mixed with radiolabeled 29FP oligonucleotides. The incubation took place at 4°C and at the indicated time points, aliquots of the mixture were taken and analyzed in EMSA.

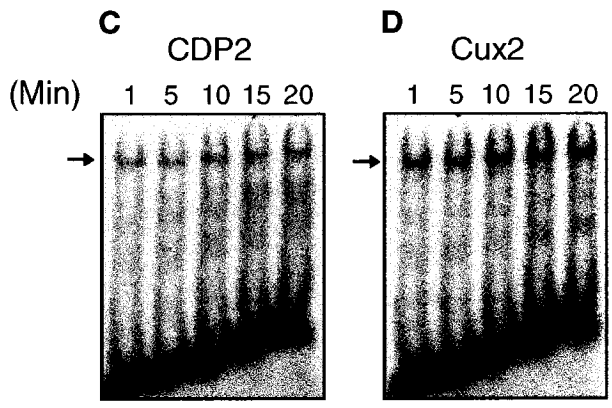
(E) Schematic representation of Myc/Cux2/HA and Myc/CDP2/HA construct used in transient transfection is presented.

OFF RATES:



EMSA: 29FP ATCG
CGAT**ATCGAT**

ON RATES:



EMSA: 29FP ATCG
CGAT**ATCGAT**

E

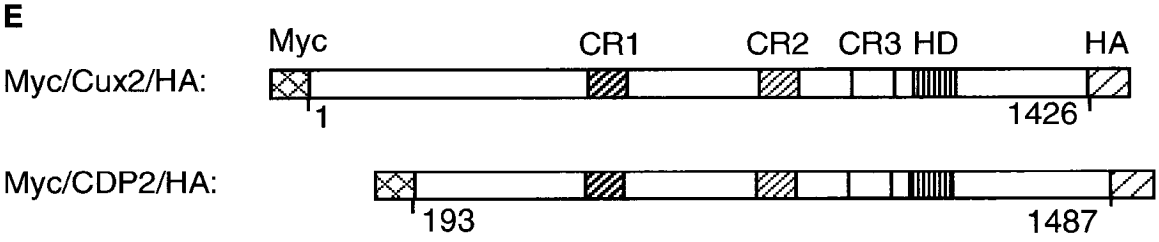


Figure 7. CDP2/Cux2 is less efficiently processed than CDP1/Cux1.

(A) NIH 3T3 cells were transfected with a vector expressing the Myc/Cux2/HA, Myc/CDP2/HA and the Myc/CDP1/HA protein. The nuclear extracts were prepared and analyzed in EMSA with oligonucleotides containing a universal CDP1/Cux1 consensus-binding site (**CGATATCGAT**). The supershifts were done with the HA Ab. The arrows show the DNA-protein complexes and the HA Ab supershift of those complexes. On the right panel, the same nuclear extracts were loaded on a 6% SDS polyacrylamide gel and transferred to PVDF membrane followed by immunoblotting using the HA Ab. The arrows show the CDP1 short and long isoforms, the CDP2 and Cux2 proteins.

(B) HS 578T cells were transfected with a vector expressing the Myc/Cux2/HA, Myc/CDP2/HA, Myc/CDP1/HA and Cux1 proteins. The nuclear extracts were prepared and loaded on a 6% SDS polyacrylamide gel and transferred to PVDF membrane followed by immunoblotting using either the HA Ab or the 1300 Cux1 Ab. The arrows show the long isoform of Cux1, CDP1, CDP2 and Cux2. The p110 short isoform of Cux1 and CDP1 is also indicated, arrow #1.

(C) Map of Cux2, indicated below the regions of the Cux2 antibodies. The expression vectors are shown.

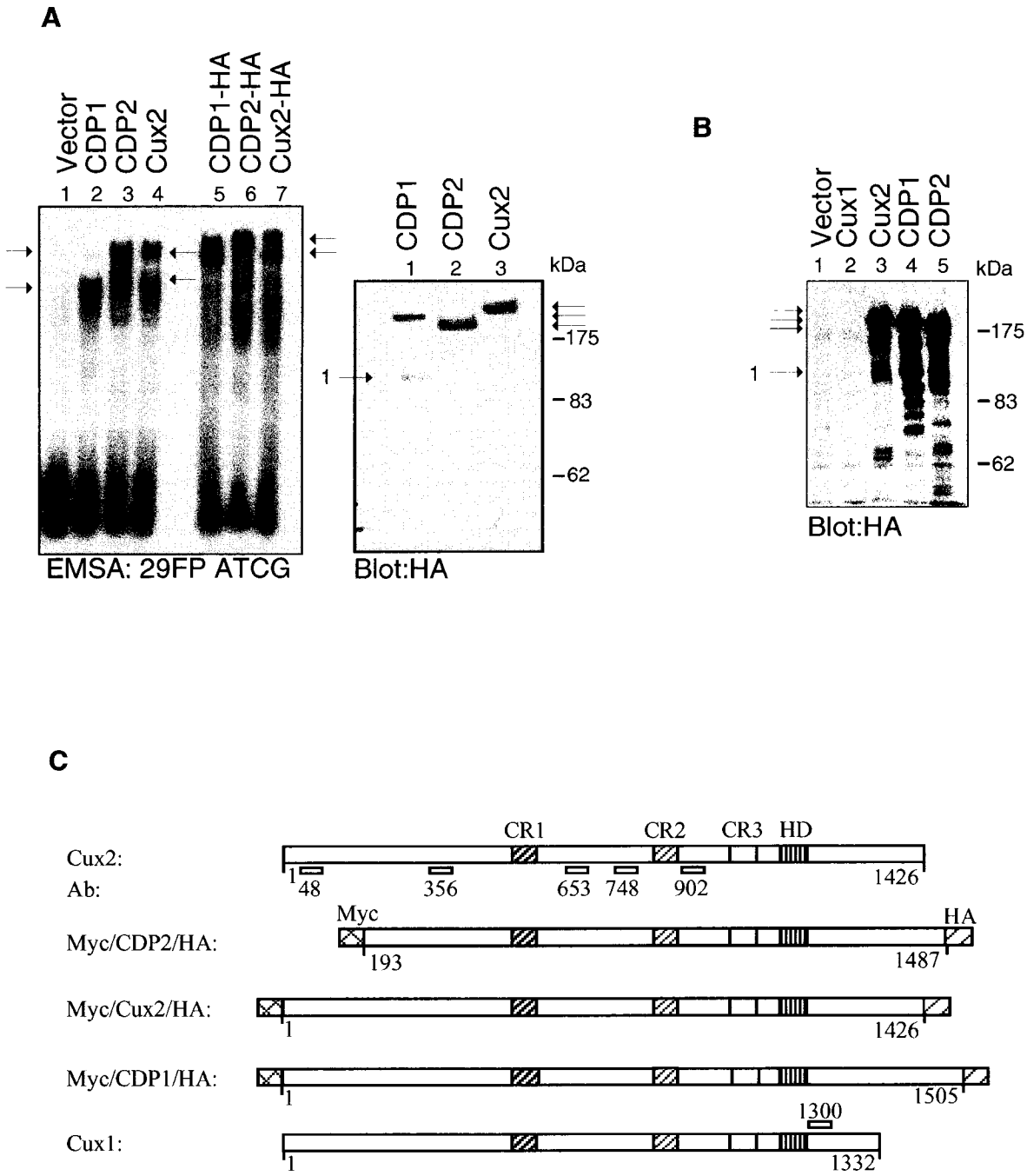
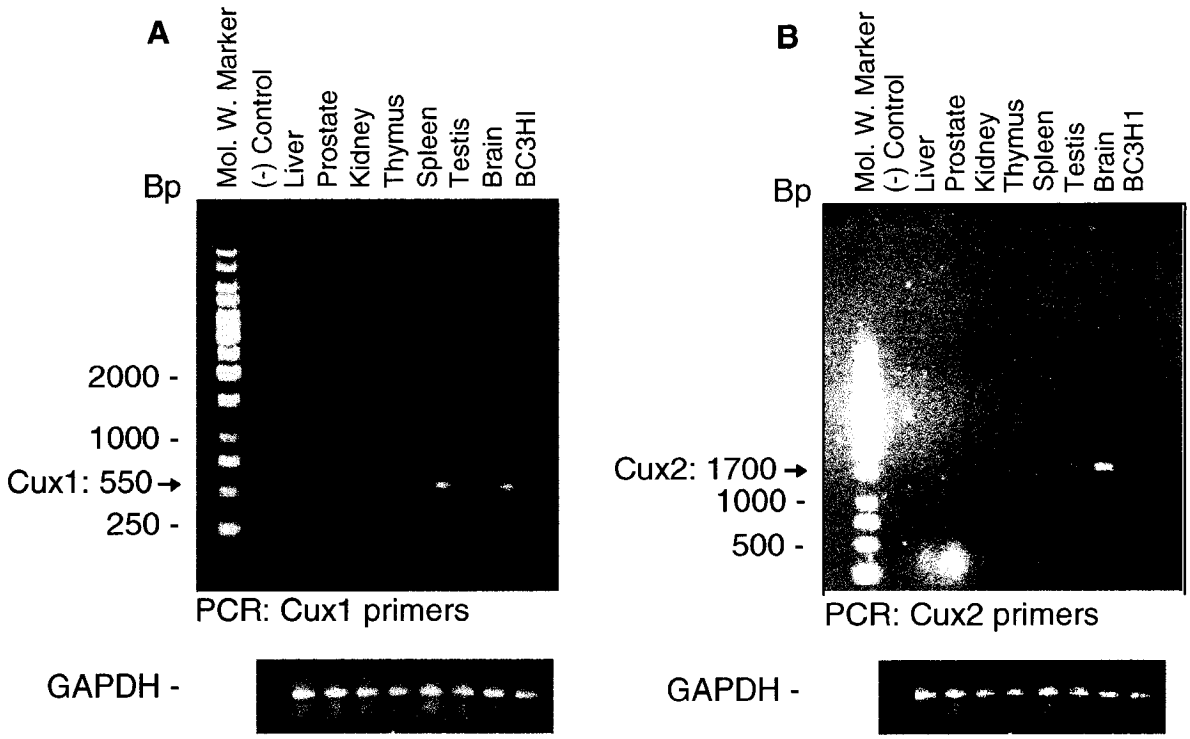

A
Vector
CDP1
CDP2
Cux2
CDP1-HA
CDP2-HA
Cux2-HA
1 2 3 4 5 6 7
EMSA: 29FP ATCG
CDP1
CDP2
Cux2
1 2 3
kDa
1
-175
-83
-62
Blot:HA
B
Vector
Cux1
Cux2
CDP1
CDP2
1 2 3 4 5
kDa
1
-175
-83
-62
Blot:HA
C
Cux2:
Ab:
CR1
CR2
CR3
HD
1
48
356
653
748
902
1426
Myc/CDP2/HA:
Myc
HA
193
1487
Myc/Cux2/HA:
1
1426
Myc/CDP1/HA:
1
1505
1300
Cux1:
1
1332

Figure 8. Cux2 mRNA is present in the brain and the testis of mice.

(A,B) RT-PCR analysis using Cux2 and Cux1 specific primers. The reverse-transcriptase reaction using the Superscript II enzyme was performed with 20 μg of total RNA isolated from different mouse tissues (See materials and methods). Then, the PCR reactions using the Taq polymerase with specific Cux2 and Cux1 primers were performed on the mouse tissues cDNA. The PCR product is 1700 bp for Cux2 and 521 bp for Cux1. 15% of the PCR product is shown. The GAPDH amplifications are shown as control for the amounts of RNA/cDNA. Mol. W. Marker: Molecular weight marker. (B) Diagrams of the Cux2 and Cux1 cDNA and the primers used for PCR amplification.


A
Mol. W. Marker
(-) Control
Liver
Prostate
Kidney
Thymus
Spleen
Testis
Brain
BC3HI
Bp
2000 -
1000 -
Cux1: 550 →
250 -
PCR: Cux1 primers
GAPDH -
B
Mol. W. Marker
(-) Control
Liver
Prostate
Kidney
Thymus
Spleen
Testis
Brain
BC3H1
Bp
Cux2: 1700 →
1000 -
500 -
PCR: Cux2 primers
GAPDH -

C

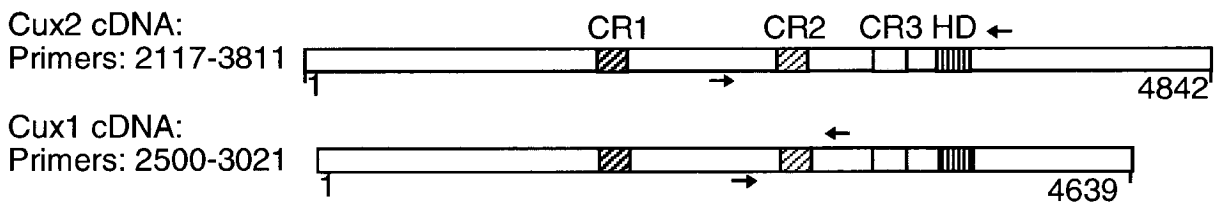
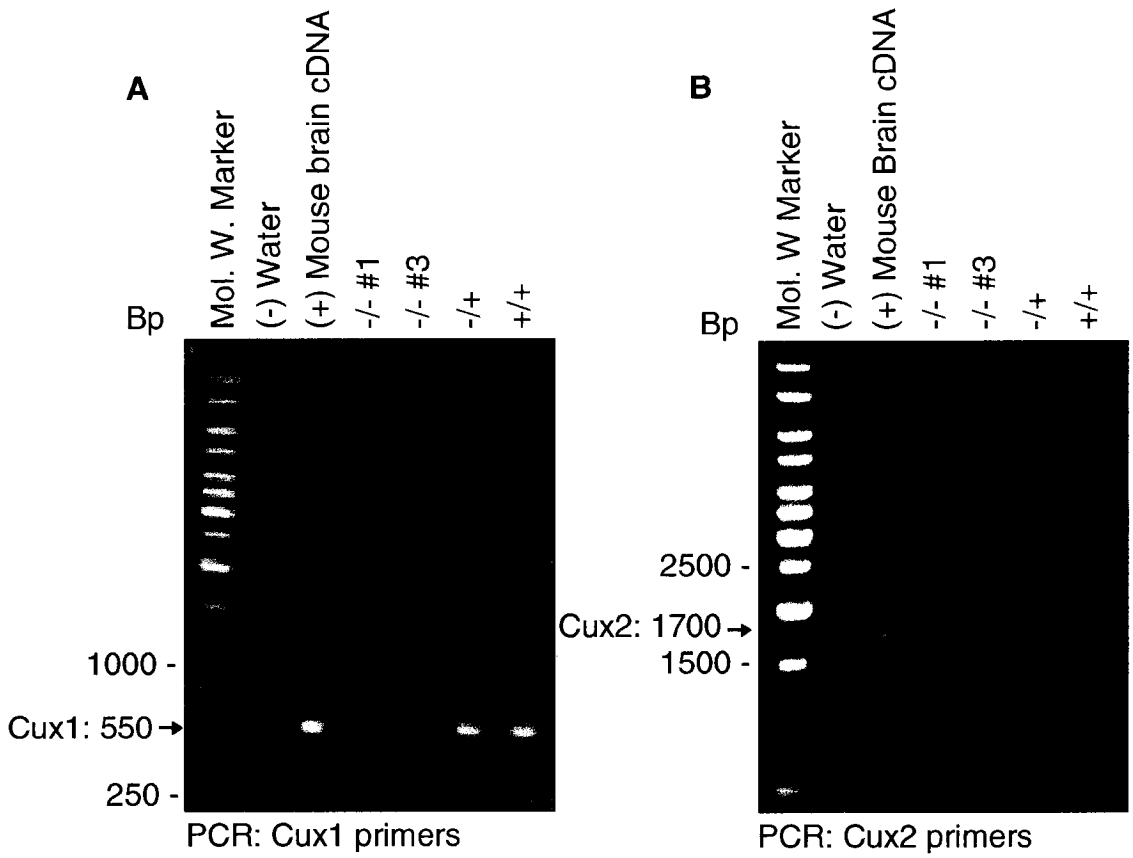

Cux2 cDNA:
Primers: 2117-3811
CR1
CR2
CR3 HD ←
1
4842
Cux1 cDNA:
Primers: 2500-3021
1
4639

Figure 9. Cux2 mRNA is not expressed in the *Cux*$^{-/-}$ mouse embryonic fibroblast (MEF) cells.

(A,B) RT-PCR analysis using Cux1 or Cux2 specific primers. MEF cells were obtained from wild type, heterozygous and homozygous null Cux1 mice, as described in materials and methods. The RT-PCR assay using the Superscript II enzyme was performed with 20 μg of RNA. PCR amplification was performed using the Taq polymerase with specific Cux2 and Cux1 primers. The PCR product is 1700 bp for Cux2 and 521 bp for Cux1. 15% of the PCR product is shown. The GAPDH amplifications are shown as control for the amounts of RNA/cDNA. The negative control is with water whereas the positive control with a mouse brain cDNA.

(C) Diagram of the Cux2 and Cux1 cDNA and the primers used for PCR amplification. Mol. W. Marker: Molecular weight marker.


A
Mol. W. Marker
(-) Water
(+) Mouse brain cDNA
-/- #1
-/- #3
-/+
+/+
Bp
1000 -
Cux1: 550
250 -
PCR: Cux1 primers
B
Mol. W Marker
(-) Water
(+) Mouse Brain cDNA
-/- #1
-/- #3
-/+
+/+
Bp
2500 -
Cux2: 1700
1500 -
PCR: Cux2 primers

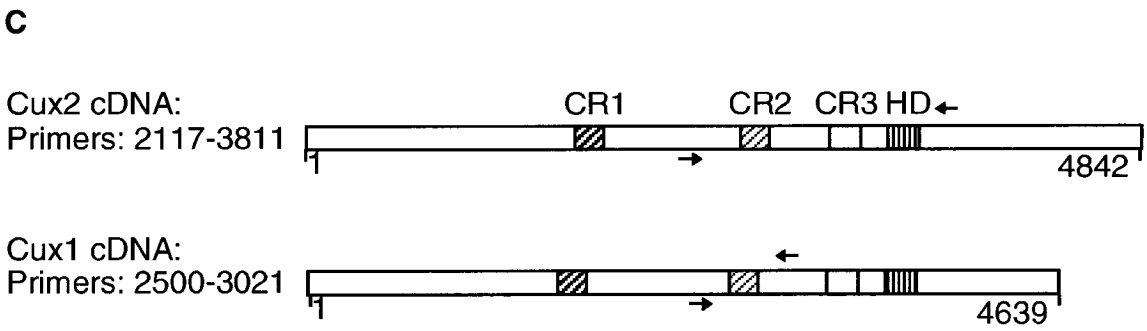
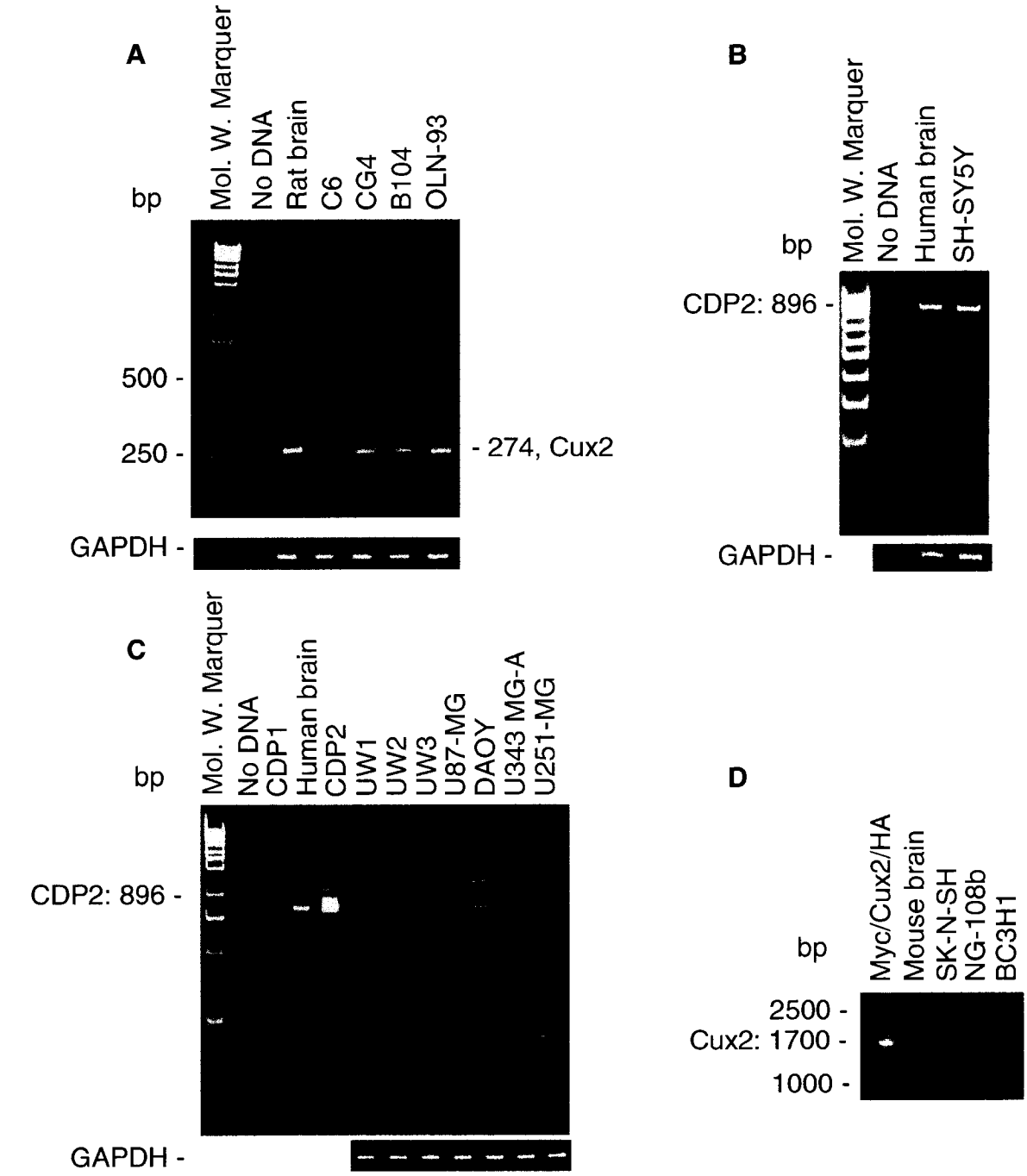

C
Cux2 cDNA:
Primers: 2117-3811
CR1
CR2
CR3
HD
1
4842
Cux1 cDNA:
Primers: 2500-3021
1
4639

Figure 10. Cux2 mRNA expression in cell lines

(A-D) RT-PCR analysis using mouse Cux2, rat Cux2 and human CDP2 specific primers. The RT-PCR assays using the Superscript II enzyme was performed with 20 μg of RNA isolated from different cell lines (See materials and methods). PCR amplification using the Taq polymerase was performed using specific mouse Cux2, rat Cux2 and human CDP2 primers. The PCR product is 1700 bp for the mouse Cux2, 174 bp for the rat Cux2 and 896 bp for the human CDP2. 15% of the PCR product is shown. The GAPDH amplifications are shown as control for the amounts of RNA/cDNA.

(E) Diagram of the mouse and rat Cux2 and human CDP2 cDNA with the primers used for the respective PCR amplification.


A
Mol. W. Marquer
No DNA
Rat brain
C6
CG4
B104
OLN-93
bp
500 -
250 -
- 274, Cux2
GAPDH -
B
Mol. W. Marquer
No DNA
Human brain
SH-SY5Y
bp
CDP2: 896 -
GAPDH -
C
Mol. W. Marquer
No DNA
CDP1
Human brain
CDP2
UW1
UW2
UW3
U87-MG
DAOY
U343 MG-A
U251-MG
bp
CDP2: 896 -
GAPDH -
D
Myc/Cux2/HA
Mouse brain
SK-N-SH
NG-108b
BC3H1
bp
2500 -
Cux2: 1700 -
1000 -

E

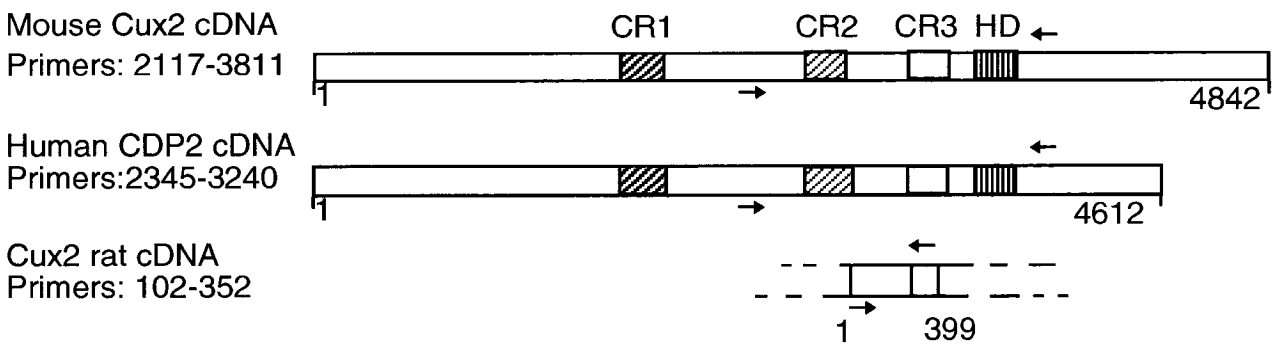

Mouse Cux2 cDNA
Primers: 2117-3811
CR1
CR2
CR3
HD
1
4842
Human CDP2 cDNA
Primers:2345-3240
1
4612
Cux2 rat cDNA
Primers: 102-352
1
399

Figure 11. Western blot analysis of Cux2 protein expression in different neuronal cell lines positive for the presence of Cux2 transcript.

(A-F) CDP1, Cux1, CDP2 and Cux2 controls lines were obtained by NIH 3T3 transfection with a vector expressing either Myc/Cux2/HA, Myc/CDP2/HA, Myc/CDP1/HA or Cux1. The nuclear extracts were prepared; 5 μg of extract was resolved on a 6% SDS polyacrylamide gel. HS 578T and FR3T3 were used as a control for the presence of CDP1/Cux1 in panel B, E, F and G. 20-30 μg of nuclear extract was loaded on the corresponding gels.

(A) The SH-SY5Y cell line was grown to 80% confluence. In parallel, the SH-SY5Y cells were treated with 40 μM of retinoic acid (RA) at 60% confluence for 3 day, to reach 80% confluence. The nuclear extracts were prepared and 50 μg of extract was resolved on a 6% SDS polyacrylamide gel and transferred to PVDF membrane followed by immunoblotting with either the 356, 748 or 861 Ab. The arrows show the CDP1 or Cux2 proteins.

(B) The N1E115 and S-KN-F1 cells were grown to 80% confluence. The S-KN-F1 cells were induced to differentiate by treating them with 100 μM of RA for 3 days until they reached 80% confluence. As controls, cells were treated with ethanol alone. The nuclear extracts were prepared and the extracts were analyzed as in (A). The nuclear pellet was resuspended in 100 μL of 1X of loading buffer (materials and methods) and 10% was loaded on the gel. The arrows show the CDP1 or CDP2 proteins.

(C) The S-KN-AS cells were grown to 80% confluence. The S-KN-AS cells were induced to differentiate by treating them with 80 μM of RA for 3 days until they reached 80% confluence. The control lanes and the nuclear pellet were prepared as in (B). The nuclear extracts were prepared and the extracts were analyzed as in (A).

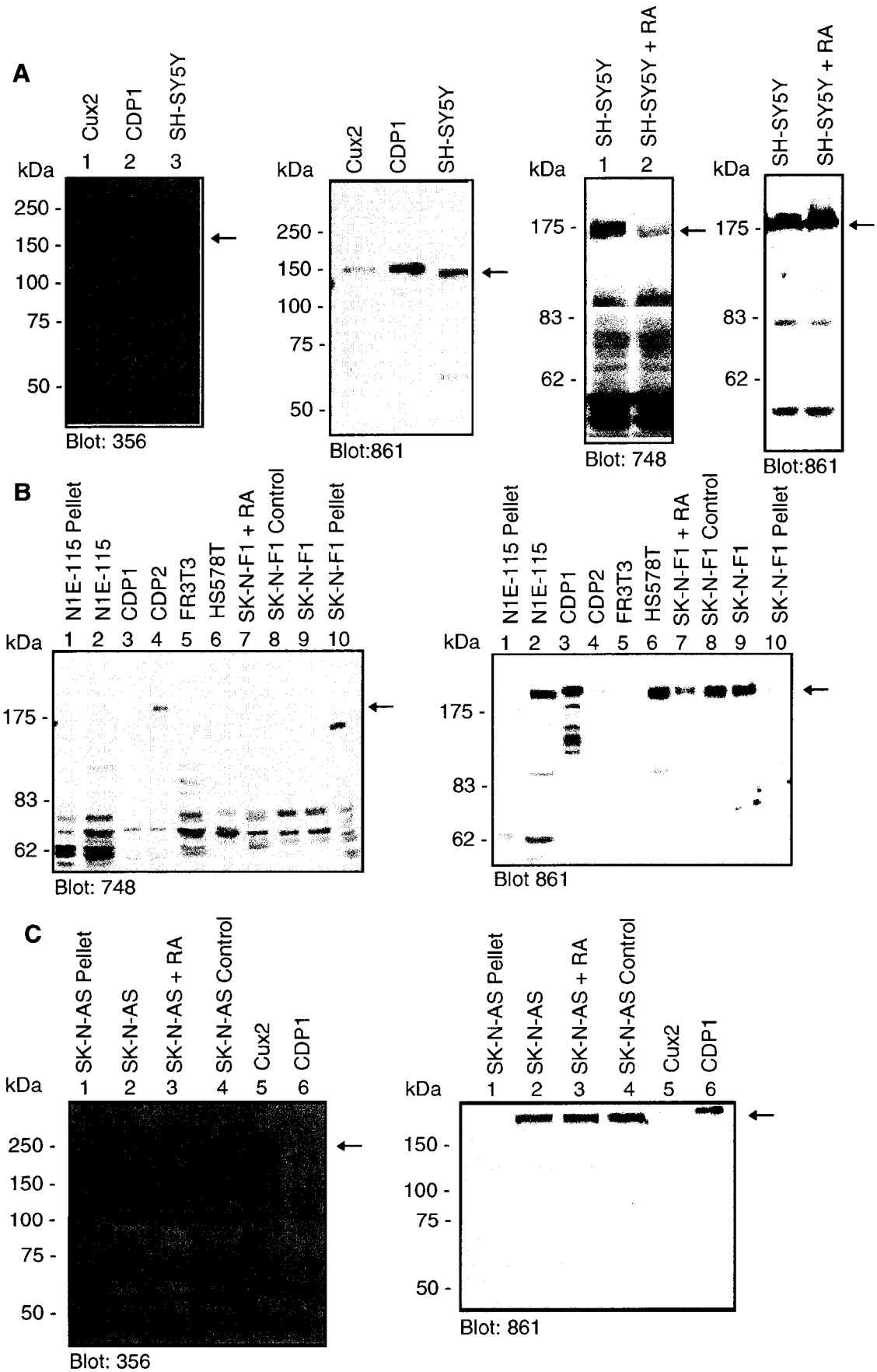
(D) N1E115 and NGP cells were grown to 80% confluence. N1E115 cells were induced to differentiate by adding DMEM without serum for 3 days. NGP cells were induced to differentiate with RA as in (C). The control wells and the nuclear pellet represent the same as in (B). The control lanes and the nuclear pellet were prepared as in (A). S-starve: Serum-starve.

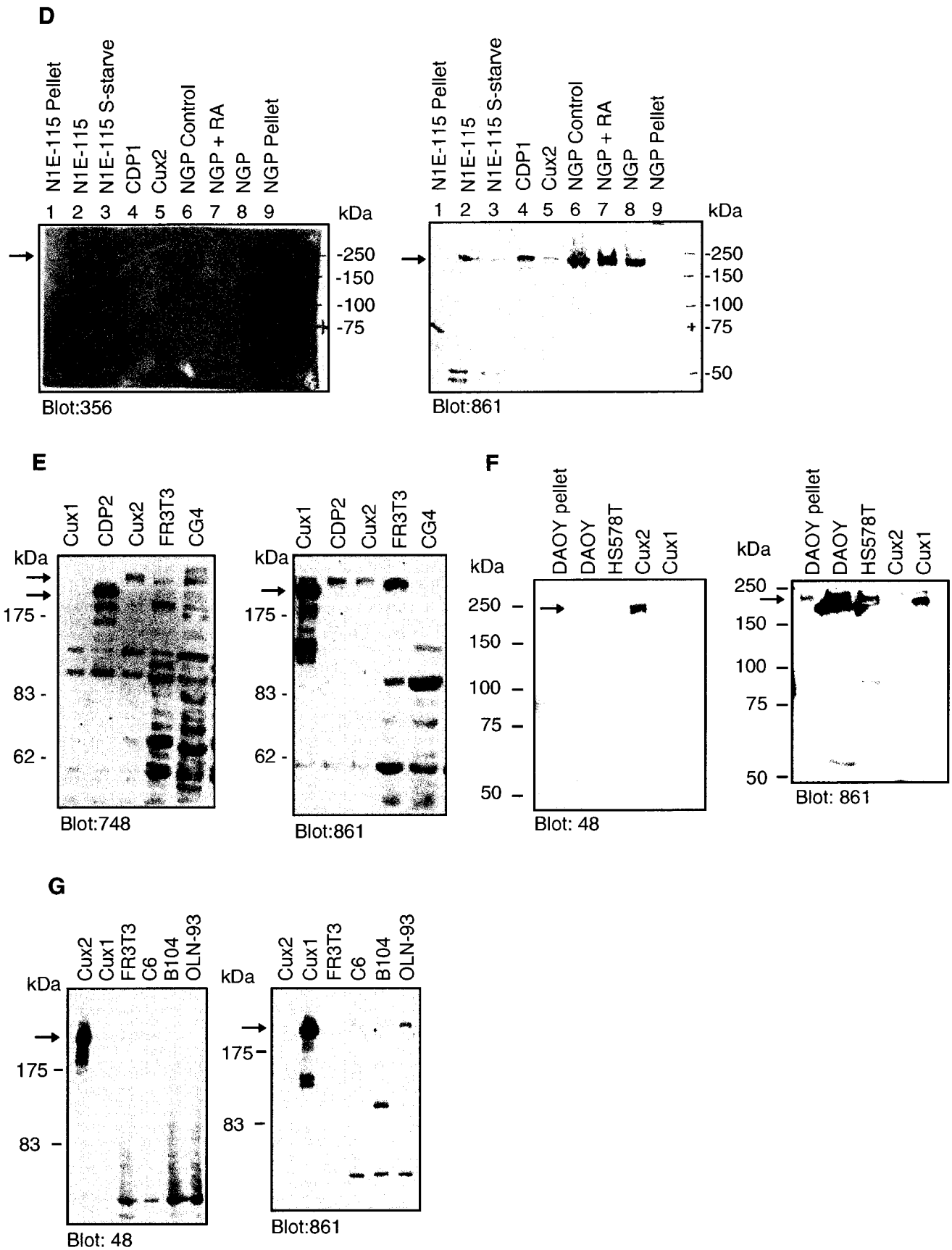
(E) The CG4 cells were grown to 80% confluence. The total extract was prepared; 100 μg of extract was resolved on a 6% SDS polyacrylamide gel. The control lanes and the nuclear pellet were prepared as in (A).

(F) DAOY cells were grown to 80% confluence. The nuclear pellet is the same as in (B). The nuclear extracts were prepared and the extracts were analyzed as in (A).

(G) C6, OLN-93 and B104 cells were grown to 80% confluence. The nuclear extracts were prepared and the extracts were analyzed as in (A).

(H) The expression vectors are shown as well as the regions recognized by each antibody.


A
Cux2
CDP1
SH-SY5Y
kDa
1
2
3
250 -
150 -
100 -
75 -
50 -
Blot: 356
Cux2
CDP1
SH-SY5Y
kDa
250 -
150 -
100 -
75 -
50 -
Blot:861
SH-SY5Y
SH-SY5Y + RA
kDa
1
2
175 -
83 -
62 -
Blot: 748
SH-SY5Y
SH-SY5Y + RA
kDa
175 -
83 -
62 -
Blot:861
B
N1E-115 Pellet
N1E-115
CDP1
CDP2
FR3T3
HS578T
SK-N-F1 + RA
SK-N-F1 Control
SK-N-F1
SK-N-F1 Pellet
kDa
1 2 3 4 5 6 7 8 9 10
175 -
83 -
62 -
Blot: 748
N1E-115 Pellet
N1E-115
CDP1
CDP2
FR3T3
HS578T
SK-N-F1 + RA
SK-N-F1 Control
SK-N-F1
SK-N-F1 Pellet
kDa
1 2 3 4 5 6 7 8 9 10
175 -
83 -
62 -
Blot 861
C
SK-N-AS Pellet
SK-N-AS
SK-N-AS + RA
SK-N-AS Control
Cux2
CDP1
kDa
1 2 3 4 5 6
250 -
150 -
100 -
75 -
50 -
Blot: 356
SK-N-AS Pellet
SK-N-AS
SK-N-AS + RA
SK-N-AS Control
Cux2
CDP1
kDa
1 2 3 4 5 6
150 -
100 -
75 -
50 -
Blot: 861


D
N1E-115 Pellet
N1E-115
N1E-115 S-starve
CDP1
Cux2
NGP Control
NGP + RA
NGP
NGP Pellet
1 2 3 4 5 6 7 8 9
kDa
-250
-150
-100
-75
-50
Blot:356
Blot:861
E
Cux1
CDP2
Cux2
FR3T3
CG4
kDa
175 -
83 -
62 -
Blot:748
Blot:861
F
DAOY pellet
DAOY
HS578T
Cux2
Cux1
kDa
250
150
100
75
50
Blot: 48
Blot: 861
G
Cux2
Cux1
FR3T3
C6
B104
OLN-93
kDa
175
83
Blot: 48
Blot:861

H

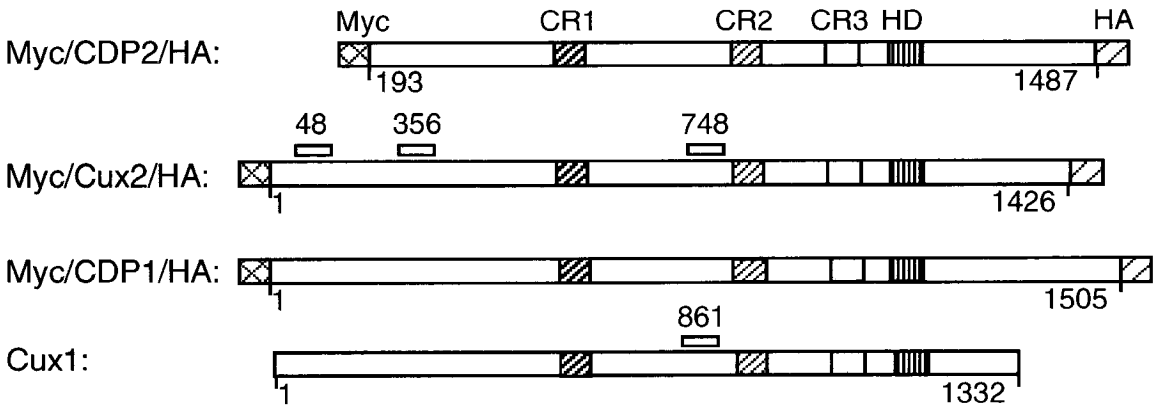

Myc/CDP2/HA:
Myc
CR1
CR2
CR3
HD
HA
193
1487
48
356
748
Myc/Cux2/HA:
1
1426
Myc/CDP1/HA:
1
1505
861
Cux1:
1
1332

Table XII. Cell lines that were tested for the presence of CDP2/Cux2 transcript or protein.

List of the cell lines that were investigated by RT-PCR and Western blot analyses. + means that CDP/Cux2 mRNA or protein were observed. Nd: not determined.

Organisms	Cell lines	Type of cells	CDP1/Cux1		CDP2/Cux2	
			RNA	Protein	RNA	Protein
Rat	OLN-93	Oligodendrocytes	+	+	+	-
	CG4	Oligodendrocytes	+	+	+	-
	B104	Neuroblastoma	+	+	+	-
	C6	Astrocytes glioma	+	+	+	-
Human	UW1	Medulloblastoma	+	nd	-	nd
	UW2	Medulloblastoma	+	nd	-	nd
	UW3	Medulloblastoma	+	nd	-	nd
	DAOY	Medulloblastoma	+	+	+	-
	SK-N-SH	Neuroblastoma	+	nd	-	nd
	SK-N-AS	Neuroblastoma	nd	+	nd	-
	SK-N-F1	Neuroblastoma	+	+	nd	-
	SY5Y	Neuroblastoma	nd	+	+	+
	NGP	Neuroblastoma	+	-	+	-
	U251-MG	Glioblastoma	+	nd	-	nd
	U343 MG-A	Glioma	+	nd	-	nd
	U87-MG	Glioblastoma-astrocytoma	+	nd	-	nd
Mouse	N1E-115	Neuroblastoma	nd	+	nd	-
	BC3HI	Mouse brain tumor	nd	nd	-	nd
Mouse/rat	NG-108B	Neuroblastoma X glioma hybrid	nd	nd	-	nd

Results

1. CDP2/Cux2 are localized in the nucleus. A putative nuclear localization signal is localized between the CR3 and the homeodomain of CDP1/Cux1 and this sequence is conserved in CDP2/Cux2. Indeed, it was shown that a full-length CDP1/Cux1 protein is localized to the nucleus, however, an amino-terminal peptide that was truncated at CR3 localized to the cytoplasm only (37). To investigate the cellular localization of CDP2 and Cux2, expression vectors were prepared with Myc and HA epitope tags at the N- and C-termini, respectively. The plasmids were introduced into NIH3T3 or HS 578T cells by transient transfections, and indirect immunofluorescence analysis was performed using HA and Myc antibodies. Nuclear staining was observed for both Cux2 and CDP2 (Fig. 1A, B). Interestingly, the two proteins displayed different signals. While CDP2 produced a relatively homogeneous signal throughout the nucleus, Cux2 produced a punctuate staining suggesting that it might be localized to specific nuclear structures or compartments. This observation was also made when the transfections were done in HS 578T cells. Since the CDP2 construct misses the first 192 N-terminal amino acids, these results suggest that the corresponding region in Cux2 is responsible for the specific nuclear sub-localization (Fig. 1C). That particular nuclear localization was not observed with CDP1/Cux1. Instead, CDP/Cux1 appears to be homogeneously distributed in the nucleus. In conclusion, CDP2/Cux2 is localized to the nucleus in transient transfections assays in both NIH3T3 and HS 578T cells. In addition, my results suggest that Cux2 might be associated with specific structures in the nucleus.

2. Cux2 antibodies do not cross-react with CDP1 or Cux1 in Western blot analysis. We wanted to generate antibodies that would recognize Cux2 and CDP2 but would not cross-react with Cux1 or CDP1. The strategy we used was to raise antibodies against fusion proteins that would contain only short regions of Cux2. Two criteria served to choose these regions: their putative antigenicity as predicted from a number of algorithms, and their relative lack of conservation with the Cux1 or CDP1 proteins. In total, 6 fusion proteins were prepared. 5 of them were injected in 10 rabbits. The peptide from residues 1177 to 1401 of Cux2 could not be isolated; this one was

discarded. We obtained antibodies that would recognize Cux2 with all 5 fusion proteins (see Materials and Methods).

The Cux2 antibodies were first characterized in immunoblotting. Cells were transfected with expression vectors coding for either Cux1, Cux2, CDP1 or CDP2. A diagram of the expression vectors is shown in Fig. 2B. Nuclear extracts were prepared, and following separation by PAGE, 7 membranes were generated for analysis with individual antibodies. Expression of Cux2, CDP1 and CDP2 was confirmed in Western blot using an anti-HA antibody since these proteins included an HA-tag at their carboxy-terminus (Fig. 2A, panel HA). Expression of Cux1 was verified using antibody 1300 (Fig. 2A, panel 1300). Antibodies #48, #356, #653, #748 and #902 were all able to recognize the mouse Cux2 protein and the human CDP2 protein (Fig. 2A). However, in the case of antibody #48, it did not recognize the human protein since the recombinant CDP2 protein is truncated at its amino terminus. Interestingly, the Cux1 and CDP1 proteins were not recognized at all by antibodies #48, #356 and #902, whereas faint signals could be detected with antibodies #653 and #748. In summary, I have generated 5 antibodies that can recognize both CDP2 and Cux2 proteins in Western blot analysis, but do not cross-react with CDP1 nor Cux1.

3. Cux2 antibodies are able to supershift CDP2/Cux2 protein-DNA complexes in Electrophoresis Mobility Shift Assays (EMSA). The Cux2 antibodies were tested in EMSA. In previous studies, the CDP1/Cux1 consensus binding site was defined as ATCGAT (2, 5, 50). Subsequent studies showed that interaction between CDP1/Cux1 and DNA involves either CR1CR2, CR2CR3HD or CR3HD (44, 88). CR1CR2 was shown to bind preferably to dimers of CGAT or CAAT, whereas CR2CR3HD and CR3HD preferred ATCGAT or ATCAAT sites (88). Therefore, 29FP ATCG oligonucleotides were used in EMSA because they contain ATCGAT an additional CGAT sequence. CDP2/Cux2 could bind and form a complex with the 29FP ATCG oligonucleotides. All antibodies were able to supershift the Cux2 protein-DNA complex (Fig. 3A lane 1 to 7) and all antibodies but #48 supershifted the CDP2 protein-DNA complex (Fig. 3B lane 1 to 9). The antibody #48 did not supershift the CDP2 complex since the recombinant CDP2 protein was truncated at its amino terminus (Fig.

3B lane 7). The protein DNA complex of CDP2 or Cux2 could not be supershifted with the Cux1 861 antibody whereas this antibody could supershift the Cux1 complex (Fig. 3A line 4 and 12, 3B line 10). The purified antibodies #48, #748 and #356 could not cross-react and supershift a Cux1 complex (Fig. 3A lane 9-11). The antibody #902 and #653 could cross-react and supershift a Cux1 complex (Fig. 3A lane 13-14). In untransfected NIH 3T3 cells, which do not express Cux2, the #356 antibody could not supershift the complex whereas the 861 could, indicating again the specificity of Cux2 antibodies for a CDP2/Cux2 protein. It is know that the NIH3T3 cells express endogenous Cux1 protein. In summary, I have generated 5 antibodies that can recognize both CDP2 and Cux2 proteins in Western blot and EMSA analyses. When purified, the antibodies can supershift specifically CDP2/Cux2 complexes.

4. Cux2 DNA binding domains function in pairs and bind to specific sequences like Cux1. It was previously demonstrated that the affinity of CR1CR2 domain for a given DNA sequence depends on the presence of two half-sites that conform the CA/GAT consensus. In contrast, CR3HD showed preference for ATCGAT (88). To address whether the DNA binding domains of Cux2 are able to function like Cux1, histidine-tagged (His-tag) CR1CR2, CR23HD and CR3HD fusion proteins were expressed in bacteria and purified by affinity chromatography, as described in Materials and Methods. Then, their DNA binding properties were investigated in EMSA using either the 29FP probe that contains the ATCGAT site plus CGGT (an imperfect CGAT), the 29FP ATCG that contains the ATCGAT site plus another CGAT site and the CCAAT-displacement activity (CDA) probe that contains the CCAAT site plus CGAT. CR1CR2 of Cux2 bound with a higher affinity to the two oligonucleotides that contain two CA/GAT half-sites (Fig. 4B-D). In contrast, CR2CR3HD and CR3HD were not affected by the presence of a second CG/AAT site and bound preferably to oligonucleotides that contained the ATCGAT sequence (Fig. 4E-J). In conclusion, this experiment demonstrates that the DNA binding domains of Cux2 can cooperate to bind to specific sequences.

5. Off and on rates of Cux2 CR1CR2, CR23HD and CR3HD. Detailed studies on

the DNA binding properties of Cux1 have revealed that the CR1CR2 domains have a fast on and off rate (88). In contrast, the on rate of CR3HD was much slower than that of CR1CR2 and the interaction of CR3HD with DNA was much more stable (88). CR23HD also made a slow and stable interaction with DNA but to a lesser extent than CR3HD (44, 88). The CR1CR2 fusion protein derived from Cux2 was able to form a complex with DNA very rapidly (Fig. 5A, D). The equilibrium of the reaction was reached in less than a minute. However, the complex was not stable as it was dissociated in less than a minute also. These properties are similar to what was observed with CR1CR2 from CDP1/Cux1 (88). However, the CR23HD and CR3HD fusion proteins derived from Cux2 did not behave at all like the corresponding regions of CDP1/Cux1. The affinity of both CR23HD and CR3HD from Cux2 did not increase with time but reached a rapid and constant equilibrium after only one minute (Fig. 5E, F). Moreover, the complexes dissociated in less than a minute for CR23HD and in less than ten minutes for the CR3HD (Fig. 5B, C). Taken together, these results show that the CR1CR2, CR2CR3HD and CR3HD domains of Cux2 rapidly form complexes with DNA, but these complexes are not stable. The properties of CR1CR2 from Cux2 are similar to the corresponding region of CDP1/Cux1. In contrast, CR2CR3HD and CR3HD of Cux2 displayed different DNA binding kinetics than the corresponding domains of CDP1/Cux1, at least with the DNA probes that were tested here. These results would suggest that CDP2/Cux2 must be able to carry the CCAAT-displacement activity, like CDP1/Cux1, but may not be as efficient in other regulatory activities that require stable association with DNA.

6. The full length CDP2/Cux2 displays DNA binding kinetics and specificity similar to that of CDP1/Cux1. The full length CDP1/Cux1 protein was found to display DNA binding activity and kinetic similar to that of its CR1CR2 domain. Similarly to its paralogous protein, the full-length CDP2/Cux2 proteins rapidly formed a complex with DNA and rapidly dissociated from it (Fig. 6). Globally, this result indicates that the full-length CDP2/Cux2 proteins exhibit rapid DNA binding kinetics.

7. Cux2 is less efficiently processed compared to Cux1. The full-length CDP1/Cux1

protein was found to be proteolytically processed at the G1/S transition of the cell cycle, thereby generating the CDP1/Cux1 p110 isoform that contains three DNA binding domains, CR2, CR3, and HD (89). Another short isoform has been identified, p75, that is encoded by an mRNA which is initiated within intron 20 (44). As a first approach to verify whether Cux2 could be proteolytically processed like CDP1/Cux1, EMSA and Western blot analysis were performed using extracts from NIH3T3 cells that were transfected with recombinant CDP2 or Cux 2 expression vectors. In EMSA, the control lane with CDP1 shows that the short isoform, p110, binds strongly to the probe as compared to the full-length isoform, p200, as it was previously described (Fig. 7A, lane 2) (89). Also, both complexes are supershifted with the HA antibody (Fig. 7A, lane 5). In the lanes with CDP2 and Cux2, we observe several retarded complexes that could be grossly divided into two groups: a group of slow-migrating species and a group of faster-migrating species (Fig. 7A, lane 3, 4). All these complexes appeared to be supershifted with the HA antibody, although interpretation of the results is difficult since the supershifts of the faster complexes seem to co-migrate with the slower complexes (Fig. 7A, lane 6, 7). I interpret the highest band as being the supershift of the slow complexes, whereas the band just below would be the supershift of the faster species. I favor this interpretation because the intensity of the faster complexes is decreased in the presence of antibody. An alternative interpretation would be that the slowest migrating species represents the supershift of the slow complex, the band just below would be the slow complex itself, and the faster migrating complex would not be supershifted by the antibodies. Thus, according to the first interpretation results from EMSA would suggest that CDP2/Cux2 can be proteolytically processed in a manner analogous to that of CDP1/Cux1. In contrast, the alternative interpretation would lead us to conclude that CDP2/Cux2 is not processed like CDP1/Cux1.

In Western blot analysis using the HA antibody directed against a C-terminal epitope, both the p200 and p110 isoforms of CDP1 but only full-length CDP2 and Cux2 were observed using extracts from transfected NIH3T3 cells (Fig. 7A, HA panel lane 2, 3). Thus, results from transient transfections into NIH3T3 suggest that CDP2/Cux2 is not processed like CDP1/Cux1. To further investigate this question, the expression vectors were introduced by transient transfection into cells of the HS 578T breast tumor

cell line, which is known to over-express proteases that cleave CDP1/Cux1. As expected, in the extracts from CDP1 transfected cells the two most intense bands were p200 and p110, but many other bands were also visible. In the extracts from Cux2 and CDP2 transfected cells, other bands were visible in addition to the full-length species but there was not a major species of the size of p110 (Fig. 7B lane 3, 5) (for example of p110 see figure 7B arrow #1). What is missing from this figure is a control to demonstrate that all the shorter CDP1/Cux1 species are generated by specific proteolysis as opposed to shear degradation of the extract. Nevertheless, the results clearly indicate that CDP2/Cux2 is not proteolytically cleaved in the same manner as CDP1/Cux1.

8. Cux2 mRNA is present in the brain and the testis of adult mice.

One of my objectives was to identify cell lines in which CDP2/Cux2 was expressed in order to study its regulation. As a first screening method, I decided to use reverse-transcriptase polymerase-chain-reaction (RT-PCR) amplification. To set up the RT-PCR conditions for CDP2/Cux2 mRNA, I used RNA isolated from the brain tissue of human, mouse and rat. Here, I present results using RNA from various tissues of the mouse. *Cux1* was shown to be expressed in most tissues of the mouse both in the adult in the embryo (3, 92, 100, 126, 131). In contrast, in situ hybridization and Northern blot analyses revealed that that the *cux2* transcript is present in a dynamic expression pattern during mouse embryogenesis whereas it is only expressed in the nervous system in the adult mouse and late in development (57, 100). From my RT-PCR experiments, the *cux1* transcript was present in most tissues tested (Fig. 8A, left panel). In contrast, the *cux2* transcript is observed only in the brain and the testis (Fig. 8B, right panel). This result confirmed that the *cux2* transcript indeed is present in the brain, as previously shown. Thus, I could use RNA from the brain as a positive control in my RT-PCR analysis.

9. Cux2 mRNA is not expressed in the mouse embryonic fibroblast *Cutl1*$^{-/-}$ knockout. The *cux1*$^{-/-}$ mice express a mutant Cux1 protein whereby the C-terminal Cut repeat 3 and homeodomain exons are replaced with an in-frame *lacZ* gene. *Cux1*$^{-/-}$

mutant mice are born at Mendelian frequency on inbred genetic backgrounds, but die shortly after birth because of retarded differentiation of the lung epithelia, a finding that indicates an essential role of Cux1 in lung maturation (37). Several other defects have been reported, yet considering that the cut null mutant is embryonic lethal in Drosophila, the fact that *cux1*$^{-/-}$ mice survive until birth was rather surprising. We considered the possibility that survival of *cux1*$^{-/-}$ mice was due to compensatory expression of *cux2*. To begin to test this possibility, expression of the *cux2* transcript was tested by RT-PCR analysis in mouse embryonic fibroblasts (MEF) derived from *cux1* mutant mice. As a control, *cux1* mRNA was found to be expressed in the wild type and heterozygous mice, but not in the homozygous null *cux1* mice (Fig. 9A). The *cux2* transcript was detected in brain RNA (+ control), but not in MEF cells from any genetic background (Fig. 9B). I conclude that, at least in MEF cells, transcription of *cux2* is not activated to compensate for the mutation in the *cux1* gene.

10. Cell lines tested for the presence of Cux2 mRNA. *Cux1* was shown to be expressed in a variety of tissue culture cell lines including NIH3T3, FR3T3, HS 578T, 293T, HL60, C2, COS, HeLa, HEL. In contrast, the *cux2* transcript was not found to be expressed in any of the cell lines that are available in the laboratory (data not shown). Since the *cux2* transcript is expressed in the nervous system (100), neuronal cell lines were analyzed for the presence of *cux2* transcript (Fig. 10A-D). Different types of neuronal cell lines were analyzed: neuroblastomas, medulloblastomas, some glioblastomas as astrocytes-like cells and oligodendrocyte-like cells and a primary oligodendrocyte cell line. As controls, I used brain RNA from the corresponding species and for human and mouse I also used RNA from transfected cells. All four rat cell lines were positive for the presence of the *cux2* transcript (Fig. 10A). Only two human cell lines were positive for the *CDP2* transcript, SH-SY5Y (Fig. 10B), and DAOY (Fig. 10C). None of the mouse cell lines were positive for the *cux2* transcript. The N1E115, S-KN-AS and S-KN-F1 were first tested for the presence of CDP2/Cux2 protein rather than the transcript. It is not known if the transcript is there, but the CDP2/Cux2 protein is not present in those cell lines (Fig. 11B-D). In conclusion, RT-PCR analysis revealed that 7 out of 19 tested neuronal cell lines tested were positive for

the presence of CDP2/Cux2 transcript: C6, OLN-93, B104, CG4, SH-SY5Y, NGP and DAOY (Fig. 10A-D and table XII).

11. Western blot analysis of Cux2 protein expression in different neuronal cell lines positive for the presence of Cux2 transcript. Western blot analysis was performed to verify expression of CDP2/Cux2 protein in cell lines that were found to express the corresponding transcript. As controls for the detection of the CDP2/Cux2 protein, I used extracts from cell transfected with a CDP2 or Cux2 expression vector. As control for the quality of the extracts, I performed Western blot analysis with an antibody, 861, that recognizes CDP1 or Cux1. Indeed, the CDP1 or Cux1 protein was detected in all tested cell lines (Fig. 11A-G). In contrast, using Cux2 specific antibodies 356 and 748, a protein of the expected size was detected in only one cell line, the SH-SY5Y neuroblastoma cell line (Fig. 11A 356 panel lane 3). We observed a decrease in the expression of CDP2 upon treatment of cells with retinoic acid (RA), which is believed to induce differentiation into a more neuronal phenotype (Fig. 11A, 748 panel lane 1 and 2).

Various treatments were attempted to stimulate expression of the CDP2/Cux2 protein in the other cell lines. Induction of differentiation, whether by serum starvation in the case of N1E115 cells or by RA treatment in the case of S-KN-AS, S-KN-F1 and NGP cells, did not lead to the expression of Cux2 or CDP2 protein (Fig. 11B-D). In addition to the morphological changes noted upon induction of differentiation, we can see that treatment S-KN-F1 cells with RA and serum starvation of N1E115 cells caused a decrease in the expression of CDP1 or Cux1 proteins (Fig. 11B, 861 panel lanes 7 and 9 and 11D, 861 panel lanes 2 and 3). However, the RA treatment did not have any effect on CDP1 protein expression in S-KN-AS and NGP cells (Fig. 11C 861 panel lane 2 and 3 and 11D 861 panel lane 7 and 8).

Various protein extraction methods were tested with similar negative results (data not shown). In addition, the cell or the nuclear pellets were sometimes loaded on gel after extensive boiling (Fig. 11B, C, E, lane 1; D lane 1 and 9). I conclude that the failure to detect the CDP2/Cux2 protein is not due to the fact that the protein localizes to particular cellular compartments.

These results reveal that although the transcript of CDP2 or Cux2 was detected in many cell lines, the protein did not appear to be expressed except in the SH-SY5Y neuroblastoma cell line.

Discussion

At the time this study was initiated, not much was known on the new mammalian *cut* homologue, *CDP2/cux2*. A first characterization in mice has established that the expression of *cux2* is restricted to the nervous system in both the developing and adult mice (100). Also, it is known that *cux2* is involved in the development of the limb bud in the chicken together with *cux1*, but in different compartments (119). To confirm the expression pattern of *cux2* in the adult mice, an RT-PCR assay, which is the most sensitive technique available for mRNA detection (30), was performed. Indeed, *cux2* was shown to be present in the brain and also in the testis (Fig. 1B). The presence of *cux2* in the testis is not surprising since it is known to be a site where most genes are expressed in one way or another. Nevertheless, a recent study in mice has revealed the presence of a dynamic expression pattern of *cux2* during early mouse embryogenesis, beginning at the stage E9.5 (57). To summarize, like *cux1*, *cux2* is activated during the early stage of development in mice and as the expression of Cux2 becomes restricted to the nervous system in the adult mice, Cux1 expression or activity seems to be restricted to proliferating cells in different tissues. Interestingly, an example of a possible dynamic evolution of the Cut genes in vertebrate can be observed in the fact that in the chick embryo, *cux1* (and not *cux2*) expression in the craniofacial primordial is analogous to that of *cux2* in the mouse (57).

We have started a collaboration with Dr Olivier Cases with the Cux2 antibodies. He is now assessing the Cux2 localization in the developing and the adult mice. Up to now, he found a nuclear localization in the superficial cortex layer and a Golgi localization in some more restrictive cell populations in the adult mice, with the antibody #48. He is now investigating the localization of Cux2 during mice development and in human tissues.

In the MEF cells of *cux1*$^{-/-}$ knockout, that lack the homeodomain and the C-terminus region, *cux2* was not expressed (Fig. 2A,B). This suggests that *cux2* did not compensate the role fulfilled by *cux1* in the context of the MEF *cux1*$^{-/-}$ cells. Furthermore, we could speculate that there is no signal that will allow the activation of the *cux2* gene in an environment where the *cux1* gene is not functional. Of course, Cux2 expression was only investigate in MEF *cux1*$^{-/-}$ cells; it is possible that *cux2* is up-

regulated in other types of cells and tissues. An approach to look at this hypothesis would be to use the in situ hybridization technique and look at the localization of *cux2* in mice embryogenesis and in the adult tissues of those *cux1*$^{-/-}$ mice. This would allow to precisely identify any changes in the expression of *cux2*. The RT-PCR technique could also be used to analyze the *cux2* expression in tissues within the newborn *cux1*$^{-/-}$ mice before they died. At another level, we could look at the Cux2 protein localization in both the developing and the adult normal or null mice for *cux1*. We could speculate that the implication of Cux2 in mouse embryonic development might explain why the *Cux1*$^{-/-}$ knockout mice did not die at an early embryonic stage, as it is the case for the *Drosophila melanogaster Cut* mutant. To what extent the two genes carry redundant or complementary functions remains to be investigated.

Cux2 has preserved the unique sequence pattern of the Cut protein family; it contains the four DNA binding domains: the Cut homeodomain and the three Cut repeats. Also, it contains a coiled-coil region, a putative nuclear localization signal between the CR3 and the HD and three other undefined homologous regions (one at the amino terminus, another at the carboxy terminus and one just downstream of the CR2). In contrast, Cux2 does not contain the C-terminal alanine- and proline-rich region, known to function as a repression domain (82). Furthermore, Cux2 contains like Cux1 sequences that match the consensus phosphorylation site of protein kinase C (PKC) and casein kinase II (CKII) in each Cut repeats. It will have to be investigated whether these kinases can phosphorylate Cux2 and modulate its DNA binding activities. Moreover, Cux2 has a sequence variation within the cyclin binding motif (Cy) that is present in Cux1 (RREL) just after the homeodomain. In Cux2, there is an insertion of a methionine between the E and L (RREML). This methionine is also conserved in the human CDP2. One can speculate that this insertion prevents interaction with cyclin-dependent kinases.

A transfection assay in NIH3T3 and HS578T followed by immunofluorescence revealed that both CDP2 and Cux2 are localized to the nucleus as it is observed for CDP1/Cux1 proteins. Compared to CDP2 which shows a homogenous nuclear localization, we can distinguish a different nuclear pattern for the Cux2 protein. Cux2 seems to be localized in specific nuclear structures or compartments. This observation

was also made when transfections were done in HS578T cells. Indeed, the specific nuclear structure or compartment signal could be present in the amino terminus region of Cux2, since the CDP2 construct lacked the first 192 N-terminal amino acid (Fig. 3C). We do not know if a non-truncated CDP2 construct would adopt the same nuclear localization. Indeed, within the missing region of CDP2 construct, the homology between CDP2 and Cux2 is 63%, if we do not add the first 42 amino acid that are homologous to CDP1/Cux1. On the other hand, if we look at the homology between the homologous *Cut* proteins within the amino terminus missing region of CDP2, we observe that beside the first 42 amino acid conserved region, there is not a high homology (23%) between CDP1 and CDP2/Cux2. It is possible that within this region, there is a specific pattern that allows Cux2 to localize to specific nuclear structures or compartments. A lot of sub-nuclear compartments have been identified. Nuclear splicing speckles are thought to be storage sites for pre-mRNA splicing factors. PML bodies are implicated in both oncogenesis and viral infection. Cajal bodies and Gems are often found paired or juxtaposed; they are now considered to be two manifestations of the same structure; they are involved in snRNP biogenesis and in the trafficking of snoRNPs and snRNPs, which are different ribonucleoproteins that move through the Cajal body en route to nucleoli or splicing speckles, respectively (39, 83, 86, 113) (refer to appendix, Fig. F). This observation was not investigated further, but to determine whether it is a true nuclear compartment localization or an artifact, we could repeat the experiment using the Cux2 antibodies and look whether we could see the same pattern. Also, we could compare the localization of recombinant Cux2 proteins that are either full-length or truncated at the N-terminus. Another approach to define the nature of this localization could be by staining the cells against proteins that are known to localize in those nuclear compartments and look if the signal overlaps with that of Cux2. Finally, Cux2 specific antibodies could be used to analyze in details the localization of the endogenous CDP2 in the SH-SY5Y cell line that appears to express the CDP2 protein.

Of all the CDP1/Cux1 antibodies tested against CDP2/Cux2, only the TF2 (raised against the whole protein) and the N-terminal antibodies were able to cross-react with CDP2 or Cux2 (data not shown). As part of my project, five novel antibodies were raised against specific antigenic Cux2 regions. Four of the five Cux2 antibodies can

recognize both CDP2 and Cux2 protein in Western blot analysis and they cross-react with neither CDP1 nor Cux1 protein: #48, #356, #653 and #748. Only the #653 and #748 showed a faint signal for CDP1/Cux1. Unfortunately, the antibody #902 did not recognize the CDP2/Cux2 protein well. In the EMSA analysis, Cux2 and CDP2 protein could bind to a probe containing the universal CDP/Cux repeats consensus binding site, 29FP ATCG that contain the DNA sequence recognized by both CR1CR2 and CR3HD fusion proteins of Cux1 (Fig. 3A lane 7, 3B lane 2). The EMSA analysis revealed that the Cux2 antibodies can supershift both CDP2 and Cux2 protein/DNA complexes and this was shown to be specific since a Cux1 antibody did not supershift the CDP2/Cux2 protein-DNA complexes; moreover, none of the purified Cux2 antibodies could supershift Cux1 protein-DNA complexes (Fig. 3A lane 9-11). The unpurified antibodies cross-reacted with Cux1 protein and supershift Cux1-containing complex. Proteins present in the serum could have elevated nonspecific interactions and lead to a supershift of the CDP1/Cux1 protein-DNA complexes. It is known that if the protein concentration is too high in EMSA, proteins can oligomerize by nonspecific interactions. As the purified Cux2 antibodies could not supershift a Cux1 protein-DNA complex, they are clearly specific to Cux2 and CDP2 protein in EMSA.

Cut proteins belong to a novel class of homeodomain proteins that exhibit the unique feature of containing multiple DNA binding domains. In previous studies, the DNA binding activity of diverse GST/Cut repeat fusion proteins were analyzed, but as GST fusion proteins can exist as dimmers it could have affected their interaction with DNA (1, 2, 5, 25, 51, 112). More recently, it was demonstrated by using histidine-tag fusion proteins that the affinity of CR1CR2 domain of Cux1 for a given DNA sequence depends on the presence of two half-sites that conform to the CA/GAT consensus.

Also, it was shown that the affinity of CR3HD does not vary significantly comparatively to CR1CR2 if one of the half-site is mutated (89). Instead, it depends on the presence of an ATNNAT sequence. From preliminary results with CDP2/Cux2, we knew that the CDP2 and Cux2 proteins could bind to the universal Cut repeats consensus binding site, the 29FP ATCG probe. This initial result suggested that the DNA binding domains of CDP/Cux 1 and 2 exhibited similar DNA binding specificities.

To address whether the DNA binding domains of Cux2 are able to function in cooperation like Cux1, monomeric histidine-tagged fusion proteins containing two or three Cux2 DNA binding domains were generated: CR1CR2, CR23HD and CR3HD. We observed that the CR1CR2 of Cux2 was able of cooperative DNA binding, with a higher affinity if the CA/GAT half-sites are intact, like Cux1 (Fig. 6B-D) (88). Moreover, the binding affinity of neither CR23HD nor CR3HD changes in regard to the presence or not of a second CA/GAT half consensus site (Fig. 6E-J). Indeed, PCR-mediated site selection revealed that the CR3HD of Cux1 preferred the ATCGAT or ATCAAT sequences, and did not bind if the CAAT sequence was not preceded by AT (50). In summary, CR1CR2 of Cux2 binds with a higher affinity to a DNA sequence containing two intact CA/GAT half-sites like Cux1; the CR3HD of Cux2 can recognize a sequence containing an ATCGAT sequence. Thus, the DNA binding domain of Cux2 can cooperate and bind to specific sequences, like what was previously found for CDP1/Cux1. To determine whether Cux2 DNA binding domains could also recognize other sequences, a PCR-mediated site selection experiment would have to be performed. It would not be surprising to find that Cux2 can recognize other sequences, since about 20% of the selected sequence of Cux1 diverged greatly from any consensus and yet represented excellent binding sites (2, 5, 50, 127).

Experiments aimed at measuring the DNA binding kinetics of the Cux2 DNA binding domains revealed that CR1CR2 made a rapid but transient interaction with DNA. Thus, CR1CR2 from CDP/Cux 1 and 2 display the same DNA binding kinetics. Surprisingly, the CR23HD and the CR3HD of Cux2 also formed a rapid complex with DNA, and the stability of the complexes was relatively weak compared to what was observed with the same domains of Cux1 (Fig. 7B,C,E,F). For Cux1, the results from methylation interference assays suggested that the differences in binding kinetics between CR1CR2 and CR3HD are due to the fact that CR1CR2 binds the second and fifth positions of ATCGAT only within the major groove, whereas CR3HD is able to make contacts within both the major and minor grooves. One possibility to explain the unstable DNA binding by Cux2 CR3HD is that the DNA binding site that was used was not optimal. This would suppose that unstable DNA binding is not an intrinsic characteristic of the protein. To verify this hypothesis, we would have to perform a

PCR-mediated site selection experiment to identify the highest affinity binding site. This experiment would reveal whether the Cux2 CR3HD prefers other sequences. Such sequences could then be used to measure again the off an on rates. In addition, a methylation interference assay should be performed with the Cux 2 CR3HD. This would allow us to know whether the CR3HD could bind stably to its own specific sequences by making contacts within both the major and minor grooves of DNA.

For the full-length Cux2 and CDP2 proteins, they displayed the same DNA binding activity as CDP1/Cux1; they showed a fast on and off rate. Globally, this result indicates that the full-length CDP2/Cux2 proteins exhibit rapid DNA binding kinetics. It is not clear in the case of CDP2/Cux2 whether this fast DNA binding activity is acquired through the DNA binding characteristics of CR1CR2 or the CR3HD domains, since they both display a fast on and off rates. On the basis of my results, it is likely that the full-length CDP2/Cux2 protein can exhibit the CCAAT displacement activity.

Interestingly, if we compare the DNA binding affinity of CDP2 and Cux2 to the 29FP ATCG probe, we can see that with the same amount of extract the CDP2 protein seemed to bind better to DNA (Fig 8A, B). Indeed, there could be an inhibitory domain within the amino terminus region that could inhibit in some way Cux2 which has a complete amino-terminus sequence. An inhibitory domain has been identified within the first 100 amino acids of CDP1/Cux1.

Several neuronal cell lines were analyzed for the presence of *CDP2/Cux2* transcript and protein (Tab. XII). Globally, only eight cell lines out of seventeen expressed the *CDP2* or *cux2* transcript and only one expressed the protein: the human neuroblastomas SH-SY5Y cell line. Of all those different types of cell lines, neuroblastomas, medulloblastomas and glial type cells, there was no preference for a particular type of cells for the expression of *cux2* transcript. At least one cell line in each category expresses the transcript. Moreover, we were able to see clearly the CDP1 or Cux1 protein in all the cell lines that expresses the CDP1 or Cux1 transcript (except for the NGP cells) (Fig. 10). These results suggest that the CDP2/Cux2 proteins are not as ubiquitous as the CDP1/Cux1 proteins. It is possible that the mRNA of Cux2 is under a tight translational regulation process. Many mechanisms can be envisaged. For instance, there could be the formation of a secondary structure in the transcript 5'UTR,

that prevents the translational machinery to gain access to the start site. There could be a modification of the cap structure. Moreover, even if only a few cognate repressors that bind to repressive elements in the 3'UTR or the 5'UTR have been identified, it is possible that the *CDP2/Cux2* transcript may be under a negative regulation from either an RNA or protein molecule. As an example, the translation of the *lin-14* gene from *C. elegans* is under the regulation of its own RNA that binds to its 3'UTR sequence (45, 71, 135, 136). Furthermore, the mouse superoxide dismutase mRNA is a good model for the involvement of a 5'UTR-bound repressor proteins (47).

Additionally, for the cell lines that could undergo differentiation, either by serum starvation or retinoic acid (RA) treatment, no expression of *CDP2/Cux2* was induced. Although the experiment was done only once, treatment of SH-SY5Y cells with RA has lead to a decrease in the level of the CDP2 protein. Additionally, no change in the protein level was observed for CDP1. Additional experiments would need to be done, but we can speculate that the RA downregulates CDP2 in the SH-SY5Y cell line and not the CDP1. Moreover, it would be interesting to look at the CDP2 level by following induction of differentiation of SH-SY5Y cells with the nerve growth factor, dibutyryl cyclic AMP or TPA that will allowed a different phenotype upon the differentiated cells (138). This would reveal whether CDP2 regulation is specific to RA or to other downstream elements activated upon the differentiation. The SH-SY5Y cell line is a good experimental model in different scientific field as apoptosis studies, drug therapy studies, Parkinson, Alzheimer's, calcium channel and so on. Interestingly, the activation of the protein kinase C (PKC) leads to the neuronal differentiation of the SH-SY5Y. Since we know that the PKC consensus phosphorylation sites are conserved in CDP2, it would be interesting to look if it could phosphorylate CDP2 in its CR domains and down-modulate its DNA binding activity.

The full-length CDP1/Cux1 p200 protein is proteolytically processed at the G1/S transition of the cell cycle, thereby generating the CDP1/Cux1 p110 isoform that contains three DNA binding domains, CR2, CR3, and HD (88). This isoform is capable of stable interaction with DNA. Furthermore, it can stimulate, directly or indirectly, the expression of a reporter containing the promoter of the DNA polymerase alpha (DNA pol α) gene (123). The question was raised whereas CDP2/Cux2 could also be cleaved

to produce a short isoform with distinct intrinsic properties. The results obtained up to now suggest that there is some evidence for some short forms of CDP2/Cux2 but it is not as obvious as in the case of CDP1/Cux1. Evidence for a short form was mostly obtained in EMSA (Fig. 12A). It was less obvious to make the difference between degradation of the extract and a precise isoform in the western blot analysis with the nuclear extract from transient transfection of either CDP2/Cux2 in NIH3T3 and HS578T. Nevertheless, in the SH-SY5Y, we could also observe some potential short isoforms (Fig. 10A). To further investigate the question, to prove that the extract are not degraded we could do a Western blot analysis against high molecular weight proteins that are not known to be process, as for example CBP and p300. Also, to distinguish between unspecific and specific bands in Western blot analysis, we could use as specific competitor the peptides against which the antibodies were raised. This would allow us to visualize which are the specific bands that could represent short isoforms. It is possible that regulation of the protein is different between the transfected one in NIH3T3 or HS578T cells and the endogenous CDP2 in the SH-SY5Y. Indeed, if we compared the lane corresponding to the SH-SY5Y and the Cux2 protein, we see less putative short forms (Fig. 10A lane 1, 3). It would be an interesting issue to verify whether processing of CDP2 is related to some specific regulatory mechanism that is present only in neuroblastomas cells. Nevertheless, overall my results do not support the notion that CDP2/Cux2 is proteolycally processed in a cell cycle dependant manner like CDP1/Cux1.

To move forward with the project, a number of avenues can be pointed out. Further characterization of SH-SY5Y cells should involve the cellular localization of the protein. Also, it would be interesting to investigate the reason for the decrease in Cux2 protein upon differentiation of SH-SY5Y cells with RA. The question whether there are CDP2/Cux2 short isoforms is an exciting point. As described above, different approaches could be used to pursue this question. A lot of neuronal cell lines were analyzed for the presence of CDP2/Cux2 protein. It would be interesting to find more cell lines that express the protein to dispose of different models that each has their positive and negative sides. Indeed, the SH-SY5Y cells are difficult to maintain in culture, they grow very slowly and they are known to be difficult to transfect.

It would be important to verify whether CDP2/Cux2 can act as a transcription factor, for example whether it would activate a reporter plasmid containing the promoter of the DNA polymerase alpha, as it was demonstrated for the p110 Cux1 short isoform. We know that Cux2 binds to a *Ncam* promoter in a concentration dependant manner and that the domains of Cux2 expression in the nervous system during early embryonic development correlates with well-characterized spatiotemporal domains of *Ncam* expression in motor neurons, dorsal root ganglia and commissural neurons (57, 100). It would be interesting to develop a method to investigate if Cux2 is implicated in the expression of *Ncam*. Also, another project could be to define CDP2/Cux2 targets by chromatin immunoprecipitation. It would be interesting to find if Cux2 could be part of some transcription units.

With the collaboration of Dr Cases, we will be able to determine the protein localization in the developing and adult mice. This will be a novel discovery since we only know about the transcript localization. The presence of a Golgi localization was quite interesting in the sense that it could be a putative CASP2 protein. In fact, there is an alternative splicing event in CDP1/Cux1 that leads to the production of another distinct protein: CASP. This one contains the first fourteen exons of CDP1/Cux1, it contain the coiled-coil region but no Cut Repeats. It could be possible that this splicing event has been duplicated too, since the chromosomal region that encompass these genes is not that big, less than 400 Kbp. So far, by different alignment strategies, no sequences were found to correspond to a CASP2 gene. Indeed, there are no EST sequences that contain a part of CDP2 and a part of a sequence homologous to CASP. To really answer the question, a meticulous analysis would need to be done by trying to join together some EST sequences that show some homologies.

Furthermore, as said before, it would be interesting to determine if the CR3HD of CDP2/Cux2 can recognize other sequences with which it could make a stable interaction. Another avenue will be to define if indeed the PKC and the CKII can phosphorylate CDP2/Cux2 and modify the DNA binding activity of the protein. A further long term experiment could be to generate the knockout of *cux2* and look at the resulting effects. If it would be possible, it would be interesting to interbreed the Cux1$^{-/-}$ and Cux2$^{-/-}$ mice and look whether the lost of the two paralogous genes would lead to

a lethal phenotype as in the case of the ancestral gene *cut,* from *Drosophila melanogaster.*

Conclusion

The CDP2/Cux2 proteins display some unique characteristics compared to its paralogous gene CDP1/Cux1. They have a different tissue expression pattern; the *cux2* transcript shows a widespread expression through the mice development and is restricted to the nervous system in the adult. CDP2/Cux2 is expressed in a very few number of neuronal cell lines. The human neuroblastoma SH-SY5Y cell line was the only cell line that expresses the CDP2 protein. The currently available evidence does not support the notion that CDP2/Cux2 is processed like CDP1/Cux1. Interestingly, the CD23HD and CR3HD domains were not able to make a stable interaction with DNA, unlike the corresponding domains of CDP1/Cux1. The full length CDP2/Cux2 protein also exhibited fast on and off rates of DNA binding activity. I conclude that CDP2/Cux2 protein can perform the CCAAT displacement activity. Together, these results allow us to postulate that the CDP2/Cux2 gene has evolved differently in some ways compared to its paralogous gene and this could lead CDP2/Cux2 to fulfill unique functions.

Figure A. Phylogenetic tree analysis of cut domains and cut class homeodomains.

(A) Neighbour-joining tree of cut domains.

(B) Neighbour joining tree of cut superclass homeodomains and COMPASS homeodomains corrected for multiple substitutions. Drosophila Antennapedia, and mouse Pax-6, Brn-1 and Lhx-3 were used as an outgroup. Figure adapted from (21).

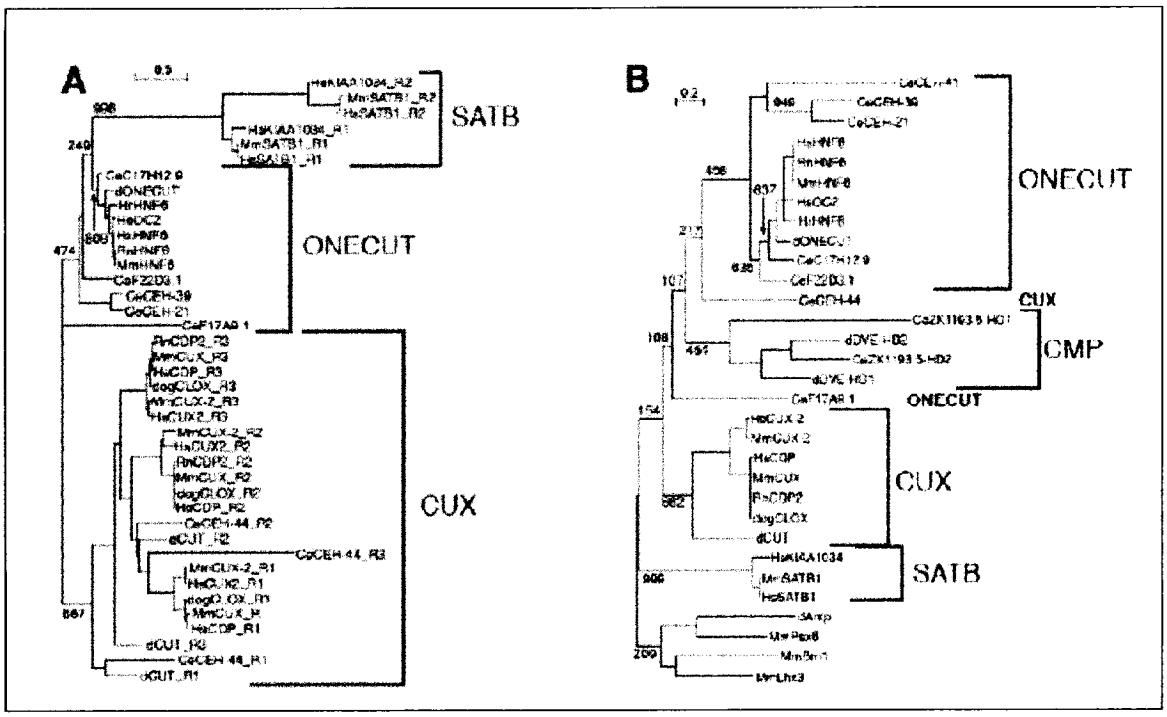

A
SATB
ONECUT
CUX
B
ONECUT
CUX
CMP
ONECUT
CUX
SATB

Figure B. Fates of duplicated genes.

(a) The classical model of degeneration of one copy after duplication.

(b) Neofunctionalisation, in which initially identical duplicates with function *A* diverge by acquiring new functions *B* and *C*.

(c) Subfunctionalisation, in which duplicate genes with multiple functions *A* and *B* diverge by reciprocal loss. The extension indicates the possibility of subsequent neofunctionalisation by acquisition of further functions *C* and *D*. Figure adapted from (84)

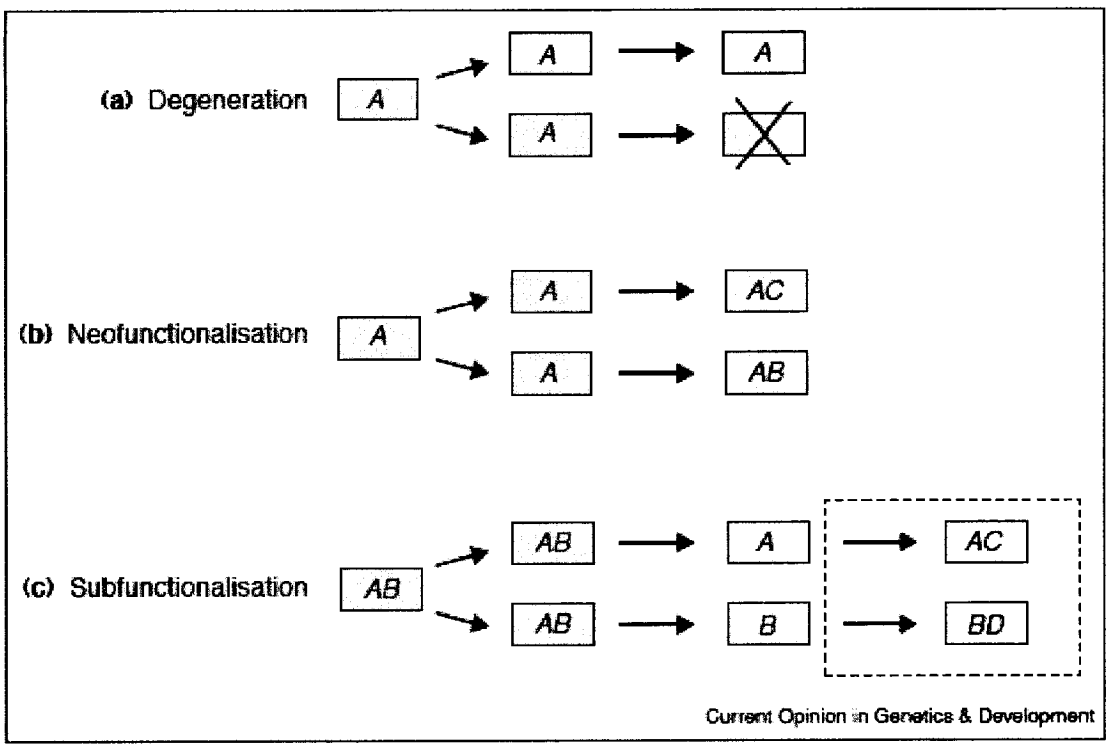

(a) Degeneration
A
A
A
(b) Neofunctionalisation
A
A
AC
A
AB
(c) Subfunctionalisation
AB
AB
A
AC
AB
B
BD
Current Opinion in Genetics & Development

Figure C. The *cut* locus in *Drosophila melanogaster*. A map of the *cut* locus is shown. White boxes represent enhancers, and the black box, the gene itself. Above the map are shown the positions of mutations, their names and the tissues affected in each case. Figure adapted from (91).

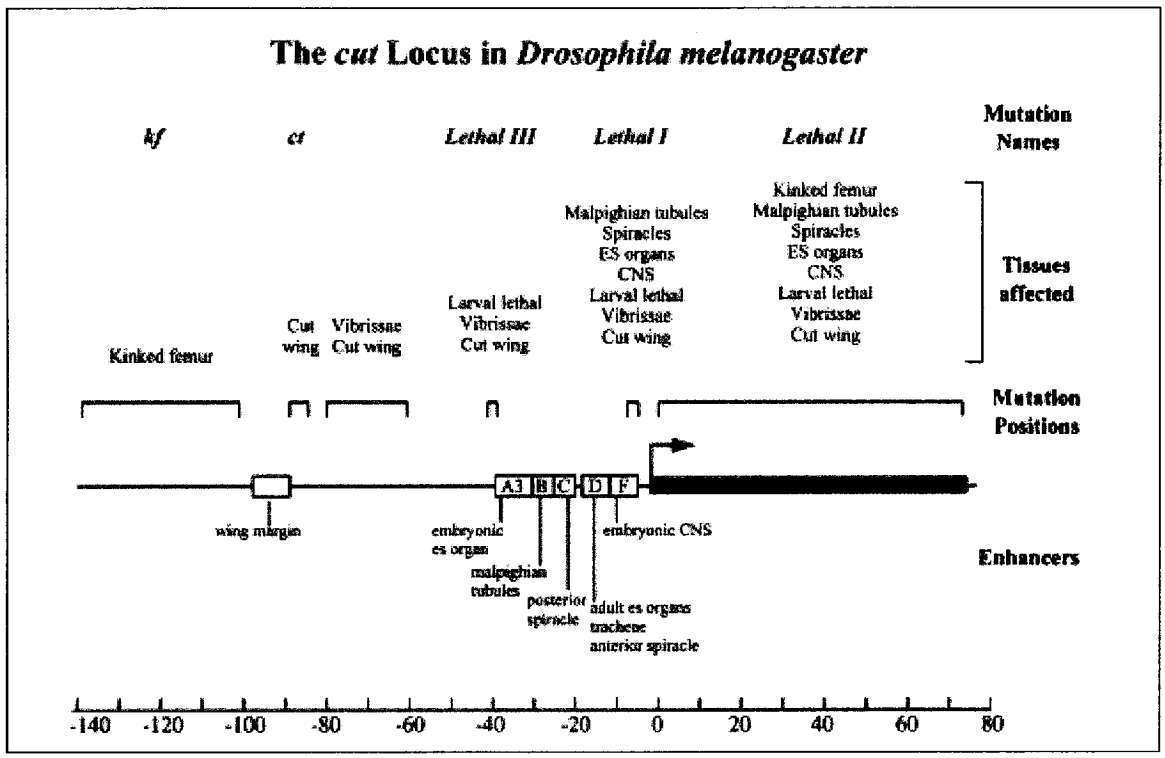

The cut Locus in Drosophila melanogaster
kf
ct
Lethal III
Lethal I
Lethal II
Mutation Names
Kinked femur
Cut wing
Vibrissae Cut wing
Larval lethal Vibrissae Cut wing
Malpighian tubules Spiracles ES organs CNS Larval lethal Vibrissae Cut wing
Kinked femur Malpighian tubules Spiracles ES organs CNS Larval lethal Vibrissae Cut wing
Tissues affected
Mutation Positions
A3
B
C
D
F
wing margin
embryonic es organ
malpighian tubules
posterior spiracle
embryonic CNS
adult es organs
tracheae
anterior spiracle
Enhancers
-140
-120
-100
-80
-60
-40
-20
0
20
40
60
80

Figure D. Expression of *Cux2* in adult brain. Bright-field (*A*) and dark-field (*B* and *C*) illumination of coronal sections of mouse brain following ISH with antisense (*A* and *B*) or sense (*C*) 33P-labeled *Cux2* riboprobes. Sections were coated with emulsion and exposed for 7 days. *3rd*, third ventricle; *lv*, lateral ventricle; *a*, arcuate hypothalamic nucleus; *am*, amygdaloid nuclei; *pc*, piriform cortex; *ec*, external capsule; *h*, hippocampus proper; *dg*, dentate gyrus; *cm*, centromedian thalamic nucleus; *ldvl*, laterodorsal thalamic nucleus (ventrolateral portion); *co*, cerebral cortex; *vp*, ventroposterior thalamic nuclei; *lpmr*, lateroposterior thalamic nucleus (mediorostral portion); *mdl*, mediodorsal thalamic nucleus (lateral portion). *Bar*, 1 mm. Figure adapted from (100).

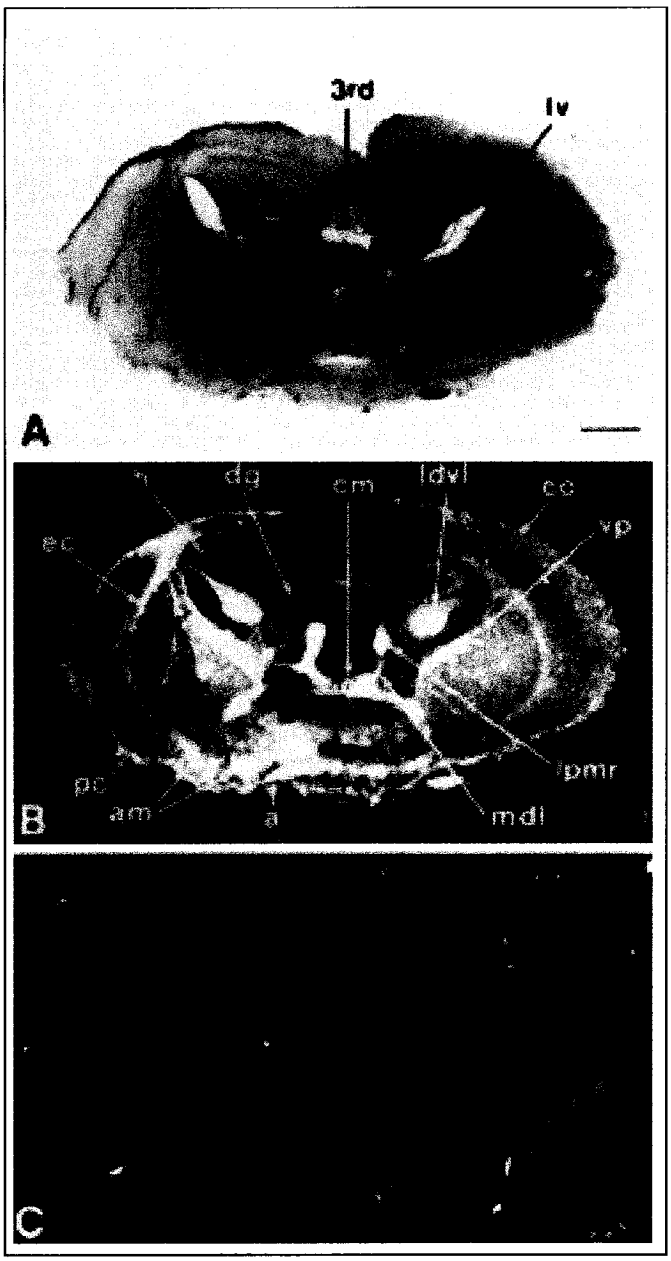
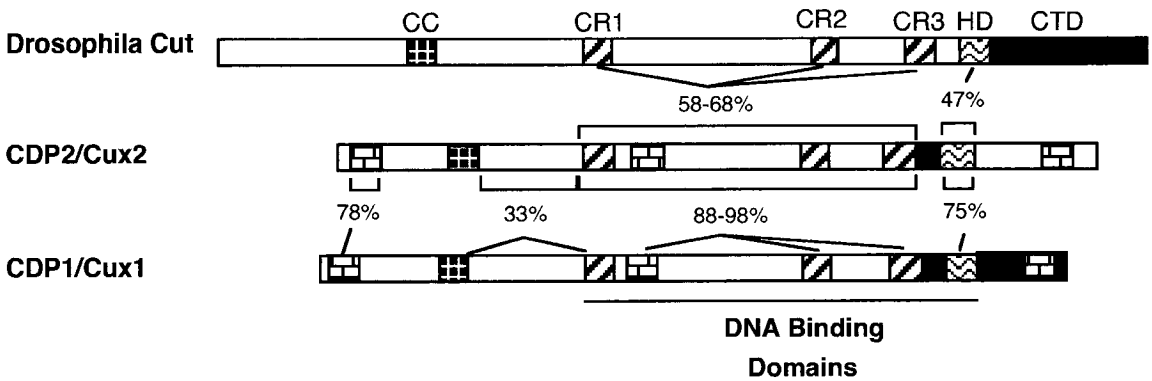

3rd
lv
A
dg
em
ldvl
cc
ec
vp
pc
am
a
mdl
pmr
B
C

Figure E. Sequence conservation between Cux2, Cux1 and Cut helps to identify functional domains. A map of the amino acids sequence of Cux2, Cux1 and Cut is shown. The percentages of homology between Cux2 and Cux1 or Cut are shown below the corresponding domain. Below the map are shown the name of the different domain.



Additional similar sequences

Coiled-coil (CC)

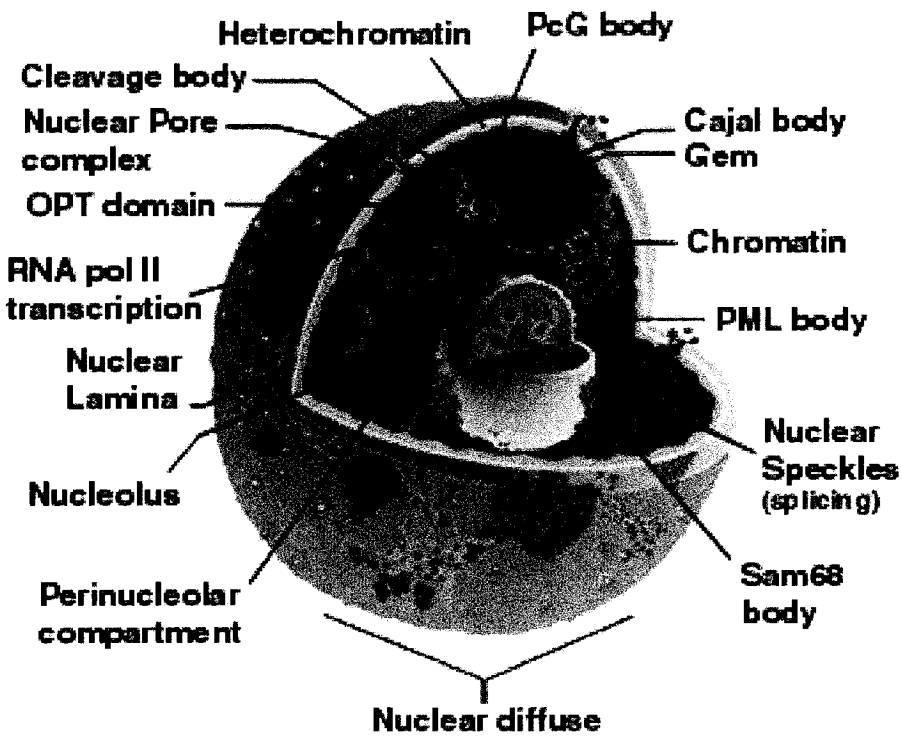
Cut Repeats (CR)

High Cationic Region (nuclear localization signal)

Homeodomain (HD)

Alanine- and Proline-rich Region; Repression domain (CTD)

Figure F. Sub-nuclear compartment of the cell. A map of the nucleus is presented. The different nuclear compartments are indicated. Figure adapted from (114).


Heterochromatin
PcG body
Cleavage body
Nuclear Pore complex
Cajal body
Gem
OPT domain
Chromatin
RNA pol II transcription
PML body
Nuclear Lamina
Nuclear Speckles (splicing)
Nucleolus
Sam68 body
Perinucleolar compartment
Nuclear diffuse

Abbreviations

#	number
^{32}P	Phosphorus-32
°C	degree celcius
aa	amino acid
Ab	antibody(ies)
AMP	Adenosine monophosphate
DH5	Escherichia coli strain
DNA	deoxyribonucleic acid
DNA pol α	DNA polymerate alpha
dNTP	Deoxyribonucleotide triphosphate
Bp	base pair
BSA	bovine serum albumine
CC	Coiled-coil
CDA	CCAAT-displacement activity
CDNA	complementary DNA
CKII	Caseine Kinase II
CO2	Carbon dioxide
CR	Cut repeat
Cy	Cyclin binding motif
DMEM	Delbecco's modified Eagle medium
CTD	carboxy-terminal domain
DTT	dithiothreitol
Dr	doctor
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetra-acetic acid
EGTA	ethylenebis(oxyethylenenitrilo)tetra-acetic acid
EMSA	electrophoresis mobility shift assay
EST	Expressed Sequence Tag
FBS	fetal bovine serum
Fig	Figure
GST	glutathione-S transferase
HA	hemagglutinin
HD	homeodomain
His	histidine
His-tag	histidine-tagged
IPTG	isopropyl-_-D-thiogaletopyranoside
Kbp	Kilobase-pair
KDa	kilodalton
MEF	Mouse embryo fibroblast
Mol. W. Marker	Molecular weight marker
min	minute
mM	mili-Molar
NCAM	nerve cell adhesion molecule
Nd	not determined

ng	nanogramme
nt	nucleotide
N-terminal	amino-terminal
PAGE	Polyacrylamide Gel Electrophoresis
PBS	phosphate buffer solution
PCR	Polymerase Chain Reaction
PKC	Proteine Kinase C
ρM	picomolar
PML	Promyelocyte Leukemia Protein
PVDF	polyvinyl di-fluoride
RA	retinoic acid
RNA	ribonucleic acid
RT	room temperature
RT-PCR	Reverse Transcriptase- Polymerase Chain Reaction
s	second
SDS	Sodium Dodecyl Sulfate
TBE	Tris-Borate-EDTA
TBS	Tris buffered saline
TBS 0,1%T	Tris buffered saline + 0,1% Tween-20
TPA	12-O-tetradecanoylphorbol-13-acetate
µl	microlitre
µg	microgramme
µM	micro-Molar
UTR	untranslated regions
v/v	volume/volume

References

1. Aceto, A., et al., *Dissociation and unfolding of Pi-class glutathione transferase. Evidence for a monomeric inactive intermediate.* Biochem J, 1992. **285**: p. 241-245.
2. Andres, V., M.D. Chiara, and V. Mahdavi, *A new bipartite DNA-binding domain: cooperative interaction between the cut repeat and homeo domain of the cut homeo proteins.* Genes & Development, 1994. **8**(2): p. 245-57.
3. Andres, V., B. Nadal-Ginard, and V. Mahdavi, *Clox, a mammalian homeobox gene related to Drosophila cut, encodes DNA-binding regulatory proteins differentially expressed during development.* Development, 1992. **116**(2): p. 321-34.
4. Artavanis-Tsakonas, S., M.D. Rand, and R.J. Lake, *Notch signaling: cell fate control and signal integration in development.* Science, 1999. **284**(5415): p. 770-6.
5. Aufiero, B., E.J. Neufeld, and S.H. Orkin, *Sequence-specific DNA binding of individual Cut repeats of the human CCAAT displacement/Cut homeodomain protein.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**: p. 7757-7761.
6. Aziz, F., et al., *HiNF-D (CDP-cut/CDC2/cyclin A/pRB-complex) influences the timing of IRF- 2-dependent cell cycle activation of human histone H4 gene transcription at the G1/S phase transition.* J Cell Physiol, 1998. **177**(3): p. 453-64.
7. Barberis, A., G. Superti-Furga, and M. Busslinger, *Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter.* Cell, 1987. **50**(3): p. 347-59.
8. Benda, P., et al., *Differentiated rat glial cell strain in tissue culture.* Science, 1968. **161**(839): p. 370-1.
9. Benos, P.V., A.S. Lapedes, and G.D. Stormo, *Is there a code for protein-DNA recognition? Probab(ilistical)ly.* Bioessays, 2002. **24**(5): p. 466-75.

10. Biedler, J.L., L. Helson, and B.A. Spengler, *Morphology and growth, tumorigenicity, and cytogenetics of human neuroblastoma cells in continuous culture.* Cancer Res, 1973. **33**(11): p. 2643-52.

11. Blanc, R., *The production of wing scalloping in Drosophila melanogaster.* Univ. Calif. Publ. Zool., 1942. **49**.

12. Blochlinger, K., et al., *Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in Drosophila.* Nature, 1988. **333**(6174): p. 629-35.

13. Blochlinger, K., et al., *Patterns of expression of cut, a protein required for external sensory organ development in wild-type and cut mutant Drosophila embryos.* Genes Dev, 1990. **4**(8): p. 1322-31.

14. Blochlinger, K., L.Y. Jan, and Y.N. Jan, *Transformation of sensory organ identity by ectopic expression of Cut in Drosophila.* Genes Dev, 1991. **5**(7): p. 1124-35.

15. Blochlinger, K., L.Y. Jan, and Y.N. Jan, *Postembryonic patterns of expression of cut, a locus regulating sensory organ identity in Drosophila.* Development, 1993. **117**(2): p. 441-50.

16. Bodmer, R., et al., *Transformation of sensory organs by mutations of the cut locus of D. melanogaster.* Cell, 1987. **51**(2): p. 293-307.

17. Bottenstein, J.E. and G.H. Sato, *Growth of a rat neuroblastoma cell line in serum-free supplemented medium.* Proc Natl Acad Sci U S A, 1979. **76**(1): p. 514-7.

18. Braun, W., *The effect of puncture on the developing wing of several mutants of Drosophila melanogaster.* J. Exp. Zool, 1942. **84**: p. 325-350.

19. Bürglin, T.R., *A comprehensive classification of homeobox genes*, in *Guidebook to the Homeobox Genes*, D. Duboule, Editor. 1994, Oxford University Press: Oxford.

20. Bürglin, T.R., *The evolution of homeobox genes*, in *Biodiversity and Evolution*, R. Arai, M. Kato, and Y. Doi, Editors. 1995, The National Science Museum Foundation: Tokyo. p. 291-336.

21. Burglin, T.R. and G. Cassata, *Loss and gain of domains during evolution of cut*

superclass homeobox genes. Int J Dev Biol, 2002. **46**(1 Spec No): p. 115-23.

22. Burgmaier, G., et al., *Association of increased bcl-2 expression with rescue from tumor necrosis factor-alpha-induced cell death in the oligodendrocyte cell line OLN-93.* J Neurochem, 2000. **75**(6): p. 2270-6.
23. Cai, H.N. and M. Levine, *The gypsy insulator can function as a promoter-specific silencer in the Drosophila embryo.* Embo J, 1997. **16**(7): p. 1732-41.
24. Castro, L. and P. Holland, *Chromosomal mapping of ANTP class homeobox genes in amphioxus: piecing together ancestral genomes.* Evol Dev, 2003. **5**(5): p. 459-465.
25. Catt, D., W. Luo, and D.G. Skalnik, *DNA-binding properties of CCAAT displacement protein cut repeats.* Cell Mol Biol (Noisy-le-grand), 1999. **45**(8): p. 1149-60.
26. Coqueret, O., G. Berube, and A. Nepveu, *DNA binding By Cut homeodomain proteins Is down-modulated by protein kinase C.* Journal of Biological Chemistry, 1996. **271**(40): p. 24862-24868.
27. Coqueret, O., G. Berube, and A. Nepveu, *The mammalian Cut homeodomain protein functions as a cell-cycle-dependent transcriptional repressor which downmodulates p21WAF1/CIP1/SDI1 in S phase.* Embo J, 1998. **17**(16): p. 4680-94.
28. Coqueret, O., et al., *DNA binding by Cut homeodomain proteins is down-modulated by casein kinase II.* Journal of Biological Chemistry, 1998. **273**(5): p. 2561-2566.
29. de Celis, J.F., A. Garcia-Bellido, and S.J. Bray, *Activation and function of Notch at the dorsal-ventral boundary of the wing imaginal disc.* Development, 1996. **122**(1): p. 359-69.
30. Dean, J., P. Goodwin, and T. Hsiang, *Comparison of Relative RT-PCR and Northern Blot Analyses to Measure Expression of ß-1,3-Glucanase in Nicotiana benthamiana Infected With Colltotrichum destructivum.* Plant Molecular Biology Reporter, 2002. **20**: p. 347-356.
31. Dickinson, L.A., C.D. Dickinson, and T. Kohwi-Shigematsu, *An atypical homeodomain in SATB1 promotes specific recognition of the key structural*

element in a matrix attachment region. J Biol Chem, 1997. **272**(17): p. 11463-70.

32. Doherty, D., et al., *Delta is a ventral to dorsal signal complementary to Serrate, another Notch ligand, in Drosophila wing formation.* Genes & Development, 1996. **10**(4): p. 421-34.
33. Dorsett, D., *Distance-independent inactivation of an enhancer by the suppressor of Hairy-wing DNA-binding protein of Drosophila.* Genetics, 1993. **134**(4): p. 1135-44.
34. Duboule, D., *Guidebook to the Homeobx Genes.* 1994, Oxford: Oxford University Press.
35. Dufort, D. and A. Nepveu, *The human cut homeodomain protein represses transcription from the c-myc promoter.* Molecular & Cellular Biology, 1994. **14**(6): p. 4251-7.
36. el-Hodiri, H.M. and M. Perry, *Interaction of the CCAAT displacement protein with shared regulatory elements required for transcription of paired histone genes.* Molecular & Cellular Biology, 1995. **15**(7): p. 3587-96.
37. Ellis, T., et al., *The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle.* Genes Dev, 2001. **15**(17): p. 2307-19.
38. Feng, X., et al., *BDNF dependence in neuroblastoma.* J Neurosci Res, 2001. **64**(4): p. 355-63.
39. Fischer, U., Q. Liu, and G. Dreyfuss, *The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis.* Cell, 1997. **90**(6): p. 1023-9.
40. Force, A., et al., *Preservation of duplicate genes by complementary, degenerative mutations.* Genetics, 1999. **151**(4): p. 1531-45.
41. Furuta, Y. and B.L. Hogan, *BMP4 is essential for lens induction in the mouse embryo.* Genes Dev, 1998. **12**(23): p. 3764-75.
42. Gaetano, C., K. Matsumoto, and C.J. Thiele, *Retinoic acid negatively regulates p34cdc2 expression during human neuroblastoma differentiation.* Cell Growth Differ, 1991. **2**(10): p. 487-93.
43. Garcia-Bustos, J., J. Heitman, and M.N. Hall, *Nuclear protein localization.*

Biochim Biophys Acta, 1991. **1071**(1): p. 83-101.

44. Goulet, B., et al., *Characterization of a Tissue-Specific CDP/Cux Isoform, p75, Activated in Breast Tumor Cells.* Cancer Research, In press: p. In press (Nov. 15, 2002).

45. Gray, N.K. and M. Wickens, *Control of translation initiation in animals.* Annu Rev Cell Dev Biol, 1998. **14**: p. 399-458.

46. Grueber, W.B., L.Y. Jan, and Y.N. Jan, *Different levels of the homeodomain protein cut regulate distinct dendrite branching patterns of Drosophila multidendritic neurons.* Cell, 2003. **112**(6): p. 805-18.

47. Gu, W. and N.R. Hecht, *Translation of a testis-specific Cu/Zn superoxide dismutase (SOD-1) mRNA is regulated by a 65-kilodalton protein which binds to its 5' untranslated region.* Mol Cell Biol, 1996. **16**(8): p. 4535-43.

48. Hamprecht, B., *Structural, electrophysiological, biochemical, and pharmacological properties of neuroblastoma-glioma cell hybrids in cell culture.* Int Rev Cytol, 1977. **49**: p. 99-170.

49. Han, M., et al., *Digit regeneration is regulated by Msx1 and BMP4 in fetal mice.* Development, 2003. **130**(21): p. 5123-32.

50. Harada, R., et al., *DNA-binding specificity of the cut repeats from the human cut-like protein.* Molecular & Cellular Biology, 1995. **15**(1): p. 129-40.

51. Harada, R., et al., *Conserved cut repeats in the human cut homeodomain protein function as DNA binding domains.* Journal of Biological Chemistry, 1994. **269**(3): p. 2062-7.

52. Heuvel, G.B.V., et al., *Expression of a Cut-related homeobox gene In developing and polycystic mouse kidney.* Kidney International, 1996. **50**(2): p. 453-461.

53. Higgy, N.A., et al., *Cux/CDP homeodomain protein binds to an enhancer in the rat c-Mos locus and represses its activity.* Biochimica et Biophysica Acta - Gene Structure & Expression, 1997. **1351**(3): p. 313-324.

54. Holland, P.W., *More genes in vertebrates?* J Struct Funct Genomics, 2003. **3**(1-4): p. 75-84.

55. Holland, P.W., et al., *Gene duplications and the origins of vertebrate*

development. Dev Suppl, 1994: p. 125-33.

56. Irvine, K.D. and E. Wieschaus, *fringe, a Boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during Drosophila wing development.* Cell, 1994. **79**(4): p. 595-606.

57. Iulianella, A., G. Vanden Heuvel, and P. Trainor, *Dynamic expression of murine Cux2 in craniofacial, limb, urogenital and neuronal primordia.* Gene Expr Patterns, 2003. **3**(5): p. 571-7.

58. Jack, J. and Y. DeLotto, *Effect of wing scalloping mutations on cut expression and sense organ differentiation in the Drosophila wing margin.* Genetics, 1992. **131**(2): p. 353-63.

59. Jack, J. and Y. DeLotto, *Structure and regulation of a complex locus: the cut gene of Drosophila.* Genetics, 1995. **139**(4): p. 1689-700.

60. Jack, J., et al., *Expression of the cut locus in the Drosophila wing margin is required for cell type specification and is regulated by a distant enhancer.* Development, 1991. **113**(3): p. 735-47.

61. Jack, J.W., *Molecular organization of the cut locus of Drosophila melanogaster.* Cell, 1985. **42**(3): p. 869-76.

62. Jackson, R.J., et al., *Human cut-like repressor protein binds TGFbeta type II receptor gene promoter.* Arch Biochem Biophys, 1999. **371**(2): p. 290-300.

63. Jacobsen, N.J., et al., *CUX2, a potential regulator of NCAM expression: genomic characterization and analysis as a positional candidate susceptibility gene for bipolar disorder.* Am J Med Genet, 2001. **105**(3): p. 295-300.

64. Kappen, C., K. Schughart, and F.H. Ruddle, *Early evolutionary origin of major homeodomain sequence classes.* Genomics, 1993. **18**(1): p. 54-70.

65. Kim, E.C., et al., *Positive and Negative Regulation Of the Human Thymidine Kinase Promoter Mediated By CCAAT Binding Transcription Factors Nf-Y/Cbf, Dbpa, and Cdp/Cut.* Cell Growth & Differentiation, 1997. **8**(12): p. 1329-1338.

66. Lannoy, V.J., et al., *Isoforms of hepatocyte nuclear factor-6 differ in DNA-binding properties, contain a bifunctional homeodomain, and define the new ONECUT class of homeodomain proteins.* J Biol Chem, 1998. **273**(22): p.

13552-62.

67. Last, T.J., et al., *Repressor elements in the coding region of the human histone H4 gene interact with the transcription factor CDP/cut.* Gene, 1998. **221**(2): p. 267-77.

68. Lawson, N.D., A. Khannagupta, and N. Berliner, *Isolation and Characterization Of the Cdna For Mouse Neutrophil Collagenase - Demonstration Of Shared Negative Regulatory Pathways For Neutrophil Secondary Granule Protein Gene Expression.* Blood, 1998. **91**(7): p. 2517-2524.

69. Ledford, A.W., et al., *Deregulated Expression of the Homeobox Gene Cux-1 in Transgenic Mice Results in Downregulation of p27(kip1) Expression during Nephrogenesis, Glomerular Abnormalities, and Multiorgan Hyperplasia.* Dev Biol, 2002. **245**(1): p. 157-71.

70. Lee, K.A., A. Bindereif, and M.R. Green, *A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing.* Gene Anal Tech, 1988. **5**(2): p. 22-31.

71. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-54.

72. Lemaigre, F.P., et al., *Hepatocyte nuclear factor 6, a transcription factor that contains a novel type of homeodomain and a single cut domain.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9460-4.

73. Li, S., et al., *Regulation of the homeodomain CCAAT displacement/cut protein function by histone acetyltransferases p300/CREB-binding protein (CBP)-associated factor and CBP.* Proc Natl Acad Sci U S A, 2000. **97**(13): p. 7166-71.

74. Li, S., et al., *Transcriptional repression of the cystic fibrosis transmembrane conductance regulator gene, mediated by CCAAT displacement protein/cut homolog, is associated with histone deacetylation.* Journal of Biological Chemistry, 1999. **274**(12): p. 7803-15.

75. Linnala, A., V.P. Lehto, and I. Virtanen, *Neuronal differentiation in SH-SY5Y*

human neuroblastoma cells induces synthesis and secretion of tenascin and upregulation of alpha(v) integrin receptors. J Neurosci Res, 1997. **49**(1): p. 53-63.

76. Liu, S. and J. Jack, *Regulatory interactions and role in cell type specification of the Malpighian tubules by the cut, Kruppel, and caudal genes of Drosophila.* Dev Biol, 1992. **150**(1): p. 133-43.

77. Liu, S., E. McLeod, and J. Jack, *Four distinct regulatory regions of the cut locus and their effect on cell type specification in Drosophila.* Genetics, 1991. **127**(1): p. 151-9.

78. Ludlow, C., R. Choy, and K. Blochlinger, *Functional Analysis Of Drosophila and Mammalian Cut Proteins In Flies.* Developmental Biology, 1996. **178**(1): p. 149-159.

79. Lundin, L.G., *Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse.* Genomics, 1993. **16**(1): p. 1-19.

80. Luo, W. and D.G. Skalnik, *CCAAT displacement protein competes with multiple transcriptional activators for binding to four sites in the proximal gp91(phox) promoter.* Journal of Biological Chemistry, 1996. **271**(30): p. 18203-18210.

81. Luong, M.X., et al., *Genetic ablation of the CDP/Cux protein C terminus results in hair cycle defects and reduced male fertility.* Mol Cell Biol, 2002. **22**(5): p. 1424-37.

82. Mailly, F., et al., *The human cut homeodomain protein can repress gene expression by two distinct mechanisms: active repression and competition for binding site occupancy.* Molecular & Cellular Biology, 1996. **16**(10): p. 5346-57.

83. Maul, G.G., et al., *Review: properties and assembly mechanisms of ND10, PML bodies, or PODs.* J Struct Biol, 2000. **129**(2-3): p. 278-87.

84. Mazet, F. and S.M. Shimeld, *Gene duplication and divergence in the early evolution of vertebrates.* Curr Opin Genet Dev, 2002. **12**(4): p. 393-6.

85. Micchelli, C.A., E.J. Rulifson, and S.S. Blair, *The function and regulation of*

cut expression on the wing margin of Drosophila: Notch, Wingless and a dominant negative role for Delta and Serrate. Development, 1997. **124**(8): p. 1485-95.

86. Mintz, P.J., et al., *Purification and biochemical characterization of interchromatin granule clusters.* Embo J, 1999. **18**(15): p. 4308-20.
87. Modolell, J., W. Bender, and M. Meselson, *Drosophila melanogaster mutations suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element.* Proc Natl Acad Sci U S A, 1983. **80**(6): p. 1678-82.
88. Moon, N.S., G. Berube, and A. Nepveu, *CCAAT displacement activity involves Cut repeats 1 and 2, not the Cut homeodomain.* Journal of Biological Chemistry, 2000. **275**(40): p. 31325-31334.
89. Moon, N.S., et al., *S Phase-Specific Proteolytic Cleavage Is Required to Activate Stable DNA Binding by the CDP/Cut Homeodomain Protein.* Molecular & Cellular Biology, 2001. **21**: p. 6332-6345.
90. Moon, N.S., et al., *Expression of N-terminally truncated isoforms of CDP/CUX is increased in human uterine leiomyomas.* Int J Cancer, 2002. **100**(4): p. 429-32.
91. Nepveu, A., *Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development.* Gene, 2001. **270**(1-2): p. 1-15.
92. Neufeld, E.J., et al., *Human CCAAT displacement protein is homologous to the Drosophila homeoprotein, cut.* Nat Genet, 1992. **1**(1): p. 50-5.
93. Neumann, C.J. and S.M. Cohen, *A hierarchy of cross-regulation involving Notch, wingless, vestigial and cut organizes the dorsal/ventral axis of the Drosophila wing.* Development, 1996. **122**(11): p. 3477-85.
94. Nirodi, C.S., et al., *The 170 kDa CCAAT displacement protein (CDP/Cut) selectively binds the IUR cis-acting element in the CXCL1 promoter : The role of CDP in the negative regulation of CXCL1 gene expression.* J Biol Chem, 2001. **22**: p. 22.
95. O'Connor, M.J., et al., *The differentiation-specific factor CDP/Cut represses transcription and replication of human papillomaviruses through a conserved*

silencing element [In Process Citation]. J Virol, 2000. **74**(1): p. 401-10.

96. Ohno, S., *Evolution by gene duplication.* 1970, Springer Verlag: Berlin.

97. Pattison, S., D.G. Skalnik, and A. Roman, *CCAAT displacement protein, a regulator of differentiation-specific gene expression, binds a negative regulatory element within the 5' end of the human papillomavirus type 6 long control region.* Journal of Virology, 1997. **71**(3): p. 2013-2022.

98. Poltorak, M., et al., *Increased neural cell adhesion molecule in the CSF of patients with mood disorder.* J Neurochem, 1996. **66**(4): p. 1532-8.

99. Popovici, C., et al., *Homeobox gene clusters and the human paralogy map.* FEBS Lett, 2001. **491**(3): p. 237-42.

100. Quaggin, S.E., et al., *Primary structure, neural-specific expression, and chromosomal localization of Cux-2, a second murine homeobox gene related to Drosophila cut.* J Biol Chem, 1996. **271**(37): p. 22624-34.

101. Richelson, E., *The use of cultured cells in the study of mood-normalizing drugs.* Pharmacol Toxicol, 1990. **66 Suppl 3**: p. 69-75.

102. Rosinski, J.A. and W.R. Atchley, *Molecular evolution of helix-turn-helix proteins.* J Mol Evol, 1999. **49**(3): p. 301-9.

103. Ross, R.A., B.A. Spengler, and J.L. Biedler, *Coordinate morphological and biochemical interconversion of human neuroblastoma cells.* J Natl Cancer Inst, 1983. **71**(4): p. 741-7.

104. Ruddle, F.H., et al., *Evolution of Hox genes.* Annu Rev Genet, 1994. **28**: p. 423-42.

105. Santaguida, M., et al., *Phosphorylation of the CCAAT Displacement Protein (CDP)/Cux Transcription Factor by Cyclin A-Cdk1 Modulates Its DNA Binding Activity in G2.* J Biol Chem, 2001. **276**(49): p. 45780-45790.

106. Saunders, J.W. and M.T. Gasseling, *Ectoderm-mesenchymal interactions in the origin of wing symmetry*, in *Epithelial Mesenchymal Interactions*, R. Fleischmajer and R. Billingham, Editors. 1968. p. 78-97:Williams and Wilkins.

107. Schubert, D., et al., *Characterization of a unique muscle cell line.* J Cell Biol, 1974. **61**(2): p. 398-413.

108. Schughart, K., C. Kappen, and F.H. Ruddle, *Duplication of large genomic*

regions during the evolution of vertebrate homeobox genes. Proc Natl Acad Sci U S A, 1989. **86**(18): p. 7067-71.

109. Scott, M.P., J.W. Tamkun, and G.W. Hartzell, *The structure and function of the homeodomain.* Biochem Biophys Acta, 1989. **989**: p. 25-48.

110. Sidow, A., *Gen(om)e duplications in the evolution of early vertebrates.* Curr Opin Genet Dev, 1996. **6**(6): p. 715-22.

111. Sinclair, A.M., et al., *Lymphoid apoptosis and myeloid hyperplasia in CCAAT displacement protein mutant mice.* Blood, 2001. **98**(13): p. 3658-3667.

112. Skalnik, D.G., E.C. Strauss, and S.H. Orkin, *CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter.* Journal of Biological Chemistry, 1991. **266**: p. 16736-16744.

113. Sleeman, J.E. and A.I. Lamond, *Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway.* Curr Biol, 1999. **9**(19): p. 1065-74.

114. Spector, D.L., *Nuclear domains.* J Cell Sci, 2001. **114**(Pt 16): p. 2891-3.

115. Spring, J., *Vertebrate evolution by interspecific hybridisation--are we polyploid?* FEBS Lett, 1997. **400**(1): p. 2-8.

116. Stunkel, W., et al., *Nuclear matrix attachment regions of human papillomavirus type 16 repress or activate the E6 promoter, depending on the physical state of the viral DNA.* Journal of Virology, 2000. **74**(6): p. 2489-501.

117. Sugimoto, T., et al., *Determination of cell surface membrane antigens common to both human neuroblastoma and leukemia-lymphoma cell lines by a panel of 38 monoclonal antibodies.* J Natl Cancer Inst, 1984. **73**(1): p. 51-7.

118. Sun, D., et al., *Enhancement of DNA ligase I level by gemcitabine in human cancer cells.* Clin Cancer Res, 2002. **8**(4): p. 1189-95.

119. Tavares, A.T., T. Tsukui, and J.C. Izpisua Belmonte, *Evidence that members of the Cut/Cux/CDP family may be involved in AER positioning and polarizing activity during chick limb development.* Development, 2000. **127**(23): p. 5133-5144.

120. Teerawatanasuk, N., D.G. Skalnik, and L.G. Carr, *CCAAT displacement protein (CDP/cut) binds a negative regulatory element in the human*

tryptophan hydroxylase gene. J Neurochem, 1999. **72**(1): p. 29-39.

121. Tontsch, U., et al., *Transplantation of an oligodendrocyte cell line leading to extensive myelination.* Proc Natl Acad Sci U S A, 1994. **91**(24): p. 11616-20.

122. Troussard, A.A., et al., *The integrin linked kinase (ILK) induces an invasive phenotype via AP-1 transcription factor-dependent upregulation of matrix metalloproteinase 9 (MMP-9).* Oncogene, 2000. **19**(48): p. 5444-52.

123. Truscott, M., et al., *CDP/Cux Stimulates Transcription from the DNA Polymerase a Gene Promoter.* Submitted to Molecular and Cellular Biology, 2003.

124. Tsugu, A., et al., *Expression of p57(KIP2) potently blocks the growth of human astrocytomas and induces cell senescence.* Am J Pathol, 2000. **157**(3): p. 919-32.

125. Tufarelli, C., et al., *Hair defects and pup loss in mice with targeted deletion of the first cut repeat domain of the Cux/CDP homeoprotein gene.* Dev Biol, 1998. **200**(1): p. 69-81.

126. Valarche, I., et al., *The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype.* Development, 1993. **119**(3): p. 881-96.

127. van den Ent, F.M., et al., *Cell cycle controlled histone H1, H3, and H4 genes share unusual arrangements of recognition motifs for HiNF-D supporting a coordinate promoter binding mechanism.* J Cell Physiol, 1994. **159**(3): p. 515-30.

128. van Gurp, M.F., et al., *The CCAAT displacement protein/cut homeodomain protein represses osteocalcin gene transcription and forms complexes with the retinoblastoma protein-related protein p107 and cyclin A.* Cancer Res, 1999. **59**(23): p. 5980-8.

129. van Wijnen, A.J., et al., *Cell cycle-dependent modifications in activities of pRb-related tumor suppressors and proliferation-specific CDP/cut homeodomain factors in murine hematopoietic progenitor cells.* Journal of Cellular Biochemistry, 1997. **66**(4): p. 512-23.

130. van Wijnen, A.J., et al., *CDP/cut is the DNA-binding subunit of histone gene*

transcription factor HiNF-D: a mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(21): p. 11516-21.

131. Vanden Heuvel, G.B., S.E. Quaggin, and P. Igarashi, *A unique variant of a homeobox gene related to Drosophila cut is expressed in mouse testis.* Biol Reprod, 1996. **55**(4): p. 731-9.

132. Vawter, M.P., et al., *Alterations of hippocampal secreted N-CAM in bipolar disorder and synaptophysin in schizophrenia.* Mol Psychiatry, 1999. **4**(5): p. 467-75.

133. von Hippel, P.H. and O.G. Berg, *Facilitated target location in biological systems.* J Biol Chem, 1989. **264**(2): p. 675-8.

134. Wang, J., et al., *Detection of hypoxia in human brain tumor xenografts using a modified comet assay.* Neoplasia, 2003. **5**(4): p. 288-96.

135. Wightman, B., et al., *Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development.* Genes Dev, 1991. **5**(10): p. 1813-24.

136. Wightman, B., I. Ha, and G. Ruvkun, *Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans.* Cell, 1993. **75**(5): p. 855-62.

137. Wu, F. and A.S. Lee, *CDP and AP-2 mediated repression mechanism of the replication-dependent hamster histone H3.2 promoter.* J Cell Biochem, 2002. **84**(4): p. 699-707.

138. Yu, V.C., et al., *Differentiation of human neuroblastoma cells: marked potentiation of prostaglandin E-stimulated accumulation of cyclic AMP by retinoic acid.* J Neurochem, 1988. **51**(6): p. 1892-9.