

**CLONING AND CHARACTERIZATION OF *RTL-1*:  
A NOVEL CUB DOMAIN-ENCODING GENE EXPRESSED IN  
THE DEVELOPING AND MATURE MOUSE NERVOUS SYSTEM**

**By**

**David Ng**

A thesis submitted in conformity with the requirements for the  
Degree of Master of Science, Graduate Department of Molecular and Medical Genetics,  
University of Toronto

© Copyright by David Ng 2001



National Library  
of Canada

Acquisitions and  
Bibliographic Services

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque nationale  
du Canada

Acquisitions et  
services bibliographiques

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file Votre référence*

*Our file Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-58701-0

**Canada**

**CLONING AND CHARACTERIZATION OF *RTL-1*: A NOVEL CUB DOMAIN-ENCODING GENE EXPRESSED IN THE DEVELOPING AND MATURE MOUSE NERVOUS SYSTEM.**

**Degree of Master of Science, 2001,**

**David Ng,**

**Graduate Department of Molecular and Medical Genetics,  
University of Toronto.**

**ABSTRACT OF THESIS**

A novel CUB domain encoding gene, *Rtl-1* (*retinal tolloid-like-1*), was identified using an *in silico* screen to find genes of potential importance to retinal development. *Rtl-1* has >40% amino acid identity with the CUB domains of tolloid and neuropilin proteins. Expression studies in mouse were initiated as a first step towards understanding the role of *Rtl-1* in mammalian neural development. Adult expression of *Rtl-1* is neuro-specific. In mature retina, *Rtl-1* is expressed specifically in bipolar cells, the major interneurons of the retina. During development, *Rtl-1* is expressed in developing neuroretina and other areas including spinal cord, cochlea, and vomeronasal organ. Two secreted isoforms (*Rtl-1A* and *Rtl-1B*) and a putative transmembrane isoform (*Rtl-1C*) are probable alternative splice products. A *Rtl-1* homolog, *Rtl-2*, will be briefly discussed. Homologs were also identified in *Caenorhabditis elegans* and *Drosophila melanogaster*. Thus, *Rtl-1* and *Rtl-2* define a novel family of putative transmembrane proteins conserved between invertebrates and vertebrates.

## ACKNOWLEDGMENTS

At last, having written the bulk of my thesis, I have finally arrived at my Acknowledgments. Although I have been looking forward to writing this section for some time, I feel I have run out of words to say. Ironically, it lies before everything else. Perhaps I should have written this part first.

I wish to thank two of my former high school teachers, Paul Edwards and Zoltan Koritar, for their enthusiasm and dedication for science. Both of you are great teachers. Also to Dr. Voon Loong (Ricky) Chan, I am grateful for your tutelage as a co-op mentor and undergraduate supervisor.

To my supervisor Dr. Roderick R. McInnes, I thank you for your guidance, encouragement, and patience throughout my studies. I sincerely appreciate the time you have taken to improve, among many things, my “mumbling” presentation style to one that is a little more audible. In addition, I appreciate the support from my committee members, Dr. Jay Cross, Dr. Joe Culotti, and Dr. Chi-Chung (C.C.) Hui.

I also would like to thank the members of the McInnes lab, both past and present, for their help, and for the countless number of unforgettable jokes and interesting conversations. To the McInnes ‘Original Nine’: Rahim Ladak, although I have been a victim of too many (of your) jokes, I can’t help but laugh at them each time you tell them again and again (and again...) to others. Thank you for your help, and for making the time fly by. Danka Vidgen, your support and conversations have been wonderful. Lenni Carreiro, Brenda Muskat, and Andrew (Andeh!!) Donnelly, the stories will last a lifetime. Jonathan Horsford, your comments have always been very helpful, constructive, and

encouraging. Thank you. I am grateful to Lynda Ploder for her excellent technical advice and warming Misha stories. Geoff Clarke and Carol Freund, I appreciate your advice, and the morning bike rides that made me wake up before eight :P

Thanks to my cousins, Janet, Shirley, Joyce, John, and Margaret, for their encouragement and thoughtfulness. Also, to Peter Harris (“Pedro”), Richard Rosenman (“Rico Sauve”), and Jeremy Letts. Thanks for being great friends. It’s amazing we’ve known each other since primary school.

Donald Wong, David McCarthy, and Toomas Karmos. Who would have thought Nebular would be what it is today?

To my parents, my warmest appreciation for your support, love, and encouragement.

Gord and Jim, thanks for being my big brothers. Jim, you have been an inspiration for me since I was a kid. Get well soon.

Lastly, I will always be in debt to Shahnaz T. Al Rashid for her advice, support, and friendship. You have shown me how to see a light when there is no light, and I thank you for that.

## TABLE OF CONTENTS

<b>Introductory Overview/Abstract of Thesis</b>	ii
<b>Acknowledgments</b>	iii
<b>Table of Contents</b>	v
<b>List of Tables</b>	vii
<b>List of Figures</b>	viii
<b>CHAPTER 1: INTRODUCTION</b>	
- Introduction	2
- The vertebrate eye and the retina	2
- Mouse eye development	6
- Morphogenesis of the mouse eye	7
- Development of the neuroretina	9
- Mouse spinal cord development	12
- Dorsoventral patterning of the spinal cord	15
- dbEST: A source for novel genes	18
- The CUB domain	20
- Tolloid and related proteins	25
- Neuropilins	28
<b>CHAPTER 2: IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF <i>Rtl-1</i></b>	
1. INTRODUCTION	32
2. MATERIALS AND METHODS	33
3. RESULTS	47
- The majority of retina-only UniGene clusters represent novel genes	47
- Identification of a novel CUB domain-encoding gene, <i>RTL-1</i>	47
- Cloning of <i>RTL-1</i>	50
- Cloning full-length mouse <i>Rtl-1</i> cDNAs	51
- Two <i>Rtl-1</i> isoforms, <i>Rtl-1A</i> and <i>Rtl-1B</i> , are putative secreted proteins	51
- Human <i>RTL-1C</i> and murine <i>Rtl-1C</i> are probable <i>Rtl-1</i> isoforms	60
- <i>Rtl-1C</i> is the most abundant splice variant in the adult mouse retina	60
- Tolloids and neuropilins share homology with <i>Rtl-1</i>	65
- <i>Rtl-1</i> has potential sites of protein modification	65
- <i>Rtl-1</i> is neurospecific	66
- Expression of <i>Rtl-1</i> in mature retinal bipolar cells	66

- <i>Rtl-1</i> is expressed in the cerebral cortex and hippocampus	69
- Expression of <i>Rtl-1</i> in the developing eye	69
- <i>Rtl-1</i> is expressed in the developing nervous system	72
- Chromosome mapping of <i>RTL-1</i> in human and mouse	72
4. DISCUSSION	75
<b>CHAPTER 3: FUTURE AND PROPOSED Ph.D. STUDIES</b>	
1. FUTURE DIRECTIONS	87
- Identification of <i>Rtl-1</i> interacting partners	87
- <i>Rtl-1</i> protein localization studies	89
- Post-translational modifications of <i>Rtl-1</i>	90
2. PROPOSED Ph.D. STUDIES	90
- Specific aims	91
- <i>Rtl-1</i> and <i>Rtl-2</i> gene targeting vectors	92
- <i>Rtl-1</i> <sup>-/-</sup> and <i>Rtl-2</i> <sup>-/-</sup> phenotypic analysis	94
- Analysis of <i>Rtl-1</i> <sup>-/-</sup> <i>Rtl-2</i> <sup>-/-</sup> double mutants	97
- In-depth analysis of the <i>Rtl-1</i> and <i>Rtl-2</i> developmental expression pattern	97
- Examination of <i>Rtl-1</i> homologs in other model organisms	98
<b>REFERENCES</b>	101

## LIST OF TABLES

Table 1-1	Summary of various CUB domain-containing proteins	23-24
Table 2-1	UniGenes representing known genes	49
Table 2-2	UniGenes representing novel genes with homology to known protein(s)	50



## LIST OF FIGURES

Figure 1-1	Structure of the vertebrate retina	4
Figure 1-2	Formation of the mouse eye	8
Figure 1-3	Formation of the vertebrate spinal cord (primary neurulation illustrated)	13
Figure 1-4	Crystallographic dimer of two aSFP spermadhesin molecules	22
Figure 1-5	Organization of various CUB domain-containing proteins	27
Figure 2-1	Amino acid alignments between Rtl-1 CUB1 and CUB2 with CUB domains of murine proteins most related to Rtl-1	48
Figure 2-2A	Amino acid and nucleotide sequences of the three <i>Rtl-1</i> splice variants. (Rtl-1A)	52-53
Figure 2-2B	<i>Rtl-1B</i> cDNA sequence	54-55
Figure 2-2C	<i>Rtl-1C</i> cDNA sequence	56-57
Figure 2-3	Differential usage of splice acceptor sites generate different Rtl-1 isoforms	58
Figure 2-4	Amino acid alignment between Rtl-1C and RTL-1C	61
Figure 2-5	Kyte-Doolittle hydropathy plot of Rtl-1C	63
Figure 2-6	Detection of different <i>Rtl-1</i> mRNA splice variants in retina by RT-PCR	64
Figure 2-7	<i>Rtl-1</i> is neurospecific	67
Figure 2-8	<i>RTL-1/Rtl-1</i> expression in the adult retina of human and mouse	68
Figure 2-9	<i>Rtl-1</i> expression in the adult mouse brain	70

Figure 2-10	Expression pattern of <i>Rtl-1</i> in the developing mouse eye	71
Figure 2-11	<i>Rtl-1</i> expression in the developing mouse spinal cord	73
Figure 2-12	Expression of <i>Rtl-1</i> in various neural regions during mouse embryogenesis	74
Figure 3-1	Gene targeting strategy to disrupt <i>Rtl-1</i> by homologous recombination	93
Figure 3-2	<i>In situ</i> hybridization of adult mouse cerebellum sections	96
Figure 3-3	Amino acid alignment between Rtl-1A and K03E5.1, a <i>C. elegans</i> predicted ORF	100

# **CHAPTER 1**

Introduction

## **INTRODUCTION**

The human brain is an organ that can, to name a few of its functions, think, remember, rationalize, dream, learn, compose music, calculate orbital trajectories, and coordinate motor movement. It is made up in part by billions of neurons forming a network of trillions of synaptic connections. Development of the mammalian brain is, put simply, complicated. How does the brain develop into such a complex organ, and will we ever understand all of the elegant molecular mechanisms involved? As Gregor Eichele poses as he ends his review on mammalian brain development research “What is perhaps the most intriguing question of all is whether the brain is powerful enough to solve the problem of its own creation” (Eichele, 1992).

To begin to understand the molecular processes involved in the development of the human brain, organisms with simpler nervous systems, such as *Caenorhabditis elegans* and *Drosophila melanogaster* (Garbor-Miklos and Rubin, 1996), are studied to understand the conserved mechanisms of neural development and differentiation. While the mouse nervous system is considerably more complex than those of invertebrates, it shares a greater resemblance to our nervous system. Several components of the mouse central nervous system (CNS), including the eye and spinal cord (Kaufman, 1992), serve as models in mammalian CNS development and will be briefly discussed.

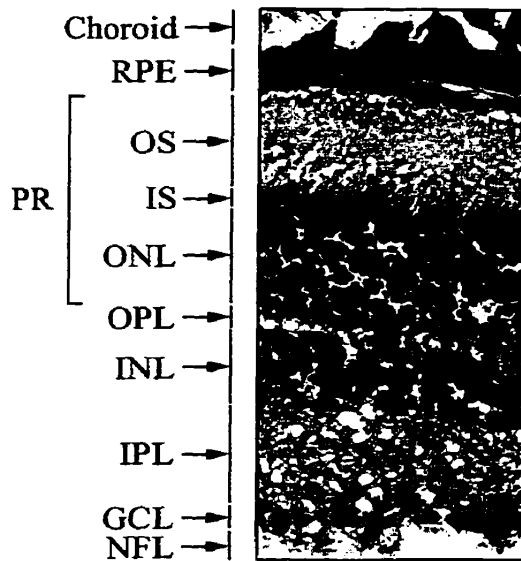
### **The vertebrate eye and retina**

The eye is an ideal model for studying the development and function of the nervous system largely because it is an accessible part of the CNS (Dowling, 1987) and yet it is a non-essential organ (Freund *et al.*, 1996). As a result, many eye mutants in

*Drosophila* and mouse exist (Oliver and Gruss, 1997; Zuker, 1994) due to the relative ease in identifying eye defects and maintaining mutant alleles. In humans, abnormalities in our vision often prompt us to quickly seek medical attention and as a result, a large number of human eye disorders have been documented. The eye is therefore well studied histologically and is relatively well characterized genetically as a large number of genes involved in eye development have been identified (Freund *et al.*, 1996).

The retina, which is situated at the back of the vertebrate eye (Rattner *et al.* 1999), is composed of the retinal pigment epithelium (RPE) and the neuroretina which together function as an integral unit. The mammalian neuroretina is easily seen to have three distinct nuclear layers called the outer nuclear layer (ONL), inner nuclear layer (INL) and ganglion cell layer (GCL) (Wechsler-Reya and Barres, 1997). Synapses between the neurons of the ONL and INL occur in the outer plexiform layer (OPL) and connections between neurons in the INL and GCL are found in the inner plexiform layer (IPL) (McIlwain, 1996). The ONL, closest to the back of the eye and immediately adjacent to the RPE, consists entirely of rod and cone photoreceptor nuclei. Rods, which contain the photopigment rhodopsin, are important for dim light (scotopic) vision. Cones mediate bright light (photopic) vision, and in some species, chromatic vision (Rattner *et al.*, 1999).

Photoreceptors are highly unique and have a specialized organelle, called the outer segment, which contains the photopigments and proteins required for phototransduction. Rod outer segments contain ~1000 stacks of flattened discs whereas cone outer segments form from a continuous stacked membrane. Photoreceptor discs are renewed daily at the base of the outer segment, while discs at the tip of the outer segment



**Figure 1-1.** Structure of the vertebrate retina.

The vertebrate neuroretina is a laminar structure and consists of three nuclear layers (shown in purple). RPE, retinal pigment epithelium; OS, outer segments; IS, inner segments; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; GCL, ganglion cell layer; NFL, nerve fiber layer (retinal ganglion cell axons); PR, photoreceptor cells. (Micrograph courtesy of Danka Vidgen).

are phagocytosed by the RPE cells (Cohen, 1970; Young, 1976; Anderson *et al.*, 1978). The synchronized renewal and turnover of discs is critical for maintaining the healthy state of the photoreceptor cells. This system has presumably evolved as a mechanism for removing proteins damaged by light (Besharse, 1986).

The inner nuclear layer (INL), consisting of 8-12 rows of densely packed nuclei, is made up of bipolar, horizontal, interplexiform and amacrine interneurons, and Müller cells, a form of glial cell. Bipolar cells are the major interneurons in the retina that transmit information from rod or cone photoreceptors to the ganglion cells (Dowling, 1987). In the absence of light (*i.e.* when unstimulated), photoreceptors, unlike most other neurons, depolarize and secrete neurotransmitters. Upon phototransduction, the neuron becomes hyperpolarized and neurotransmitter secretion is inhibited (McIlwain, 1996). Bipolar cells, which respond to the presence or absence of photoreceptor neurotransmitters (such as glutamate), are of two distinct types. The first type, known as ON bipolar cells, depolarize in the absence of neurotransmitters from photoreceptors (upon light stimulation) (Weblin and Dowling, 1969), and are important for transmitting information about bright fields of view. OFF bipolar cells, by comparison, depolarize in the presence of photoreceptor neurotransmitters (*i.e.* in the absence of light) (Kaneko, 1970) and therefore transmit information regarding dark fields of view.

Horizontal cells mediate lateral communication between photoreceptors and bipolar cells, forming synapses in the OPL. Amacrine cells are axonless and mediate lateral communication between bipolar and ganglion cells, forming synapses at the IPL. Interplexiform cells contact amacrine cells, bipolar cells, and horizontal cells, sending synaptic processes to both IPL and OPL (Dowling, 1987; McIlwain, 1996). The nuclei of

each cell type in the INL are generally present in specific locations, which are conserved between species. For example, the relatively rare horizontal cell nuclei are located towards the outer-most region of the INL (closest to the OPL) while the more abundant bipolar cell nuclei are found in the outer-half of the INL. Amacrine and interplexiform cell nuclei are situated towards the inner-half of the INL (near the IPL) (Dowling, 1987; Young, 1985).

The retinal ganglion cells (RGCs), which form the third nuclear layer, are responsible for carrying information from the neuroretina to the brain. Their axons extend along the optic nerve to the lateral geniculate nucleus (LGN) and superior colliculus in the midbrain (Dowling, 1987).

### **Mouse eye development**

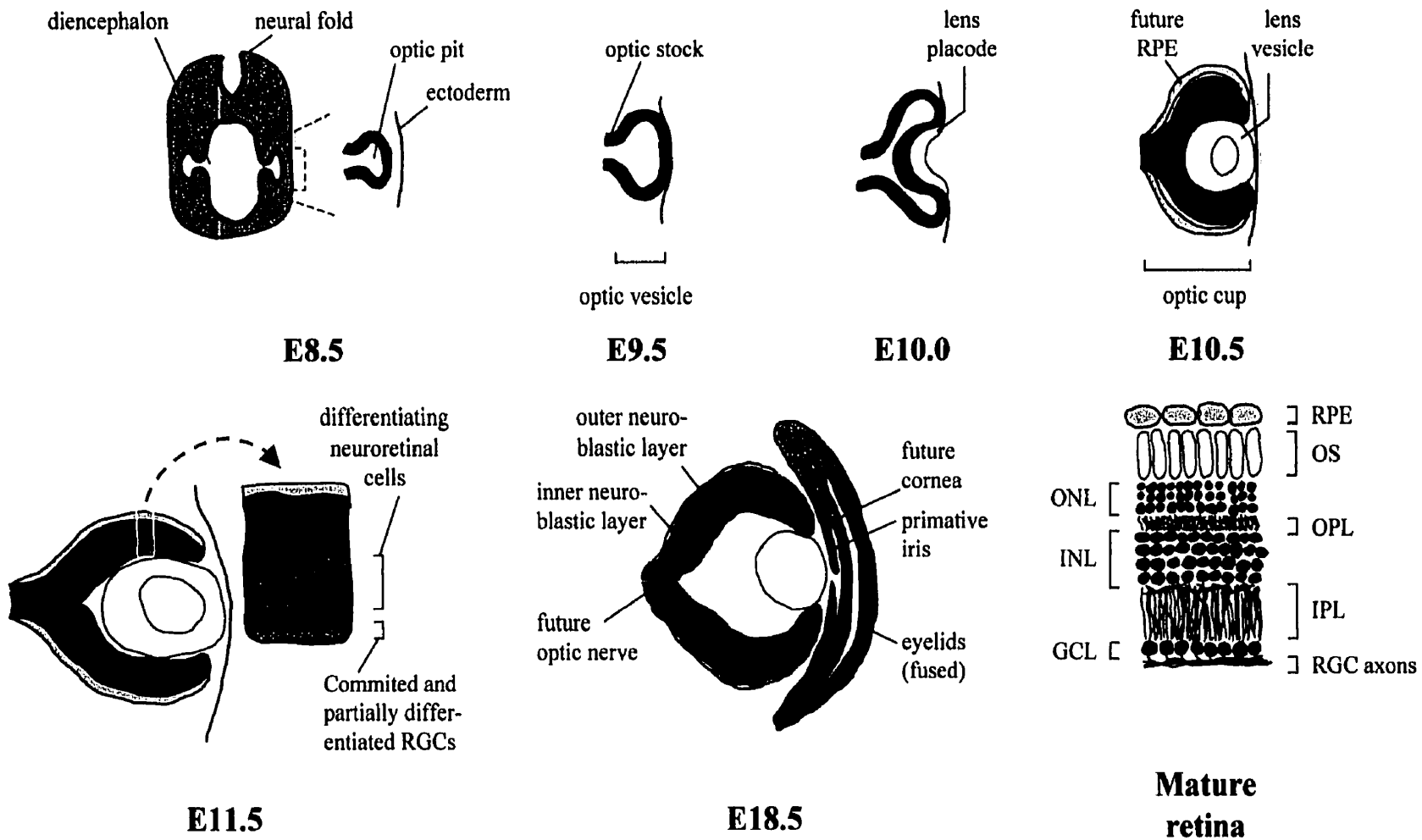
The morphological events that underlie neuroretinal development have been well documented, but the molecular mechanisms that mediate this elegant process have been largely a mystery. In the past decade, some exciting discoveries, such as functional conservation between the *Drosophila* *eyeless* and mouse Pax6 homeodomain proteins (Halder *et al.*, 1995), have provided convincing evidence that the molecular controls and genetic pathways of eye development are evolutionarily preserved. This finding was very surprising in light of the large morphological divergence in eye structures between invertebrates and vertebrates (Kumar and Moses, 1997). Because the development of the mouse eye has been extensively studied and serves as a model for mammalian eye development, a brief overview of mouse eye and retinal development is presented below.



## **Morphogenesis of the mouse eye**

The first physical sign of murine eye development begins at about E8.0 with the thickening of parts of the neural ectoderm to form the two optic placodes (Kaufmann, 1992). Between E8.5-9.0, each optic placode indents to form an optic pit, and part of the diencephalon begins to evaginate to form the optic vesicle. The optic vesicle, connected to the diencephalon by the optic stalk, contacts the surface ectoderm causing part of the ectoderm to thicken. The thickened ectoderm eventually forms the lens placode. By E10.0, the lens placode becomes indented below the surface of the ectoderm, initially forming the lens pit and later fuses to form the lens beginning at E10.5. The optic vesicle invaginates (at E10.5) to form the optic cup “trapping” the lens vesicle inside the optic vesicle. The optic cup is now composed of two layers. The cells in the outer layer of the optic cup (posterior portion) differentiate to form the RPE, while the inner portion of the optic cup containing multi-potential neuroprogenitor cells. At about E18.5, the neuroprogenitor cells can be readily distinguished into an outer and inner neuroblastic layer (the inner neuroblastic layer is lighter in appearance). The outer neuroblastic layer consists mainly of ventricular cells (neuroblasts) while the inner neuroblastic layer, as the neuroretina develops, contains an increasing number of committed and partially differentiating neurons. Additional differentiation and migration of retinal neurons occurs, which ultimately give rise to the different cell types that will eventually form the three stratified nuclear layers.

Epithelial cells in the posterior wall of the lens vesicle multiply, lose their nuclei, and elongate into transparent lens fibers (McIlwain, 1996), which fill the lumen of the vesicle (around E13.0). Mesodermal cells later fill the space between the lens vesicle



**Figure 1-2.** Formation of the mouse eye.

RPE, retinal pigment epithelium; OS, outer segments; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; GCL, ganglion cell layer; RGC, retinal ganglion cell.  
 (Figures adapted from Kaufman, 1992; Freund *et al.*, 1996)

and the surface ectoderm, and later forms the cornea. The choroid and the sclera differentiate from mesenchyme surrounding the optic cup. Smooth muscle for the iris and ciliary body form, and the sclera becomes a thick layer of dense connective tissue. (Pei and Rhodin, 1970; Kaufmann, 1992).

### **Development of the neuroretina**

The neuroblastic layer consists of mitotic neuroprogenitor cells, each of which has the potential to differentiate into any of the six major classes of retinal neurons (ganglion, bipolar, amacrine, horizontal, interplexiform, and photoreceptor cells) or Müller cells (Cepko *et al.*, 1996). These populations of cells are “born” over a period of days to weeks. (A cell’s birthday refers to when the cell has undergone its last mitotic cell division. There may be no evidence of differentiation at this time). In the developing mouse eye, the birthdates of each population of cells in the neuroretina have been extensively studied (Young, 1985).

At about E10.5, the first multipotential neuroretinal progenitor cells exit their cell cycle and migrate out of the neuroblastic layer towards the vitreal portion of the presumptive neuroretinal layer, and begin to differentiate into retinal ganglion cells (Young, 1985). The molecular determinants of cell cycle withdrawal of progenitor cells in the developing CNS are largely unknown. Recent studies of p27<sup>Kip1</sup> (a cyclin-dependant kinase inhibitor protein expressed throughout the CNS) in the mammalian retina, suggest that this protein regulates the number of late multipotent progenitor cells that exit their mitotic cycle and begin to differentiate. The expression pattern of p27<sup>Kip1</sup> in the developing neuroretina coincides with the onset of differentiation for most of the

retinal cell types (Levine *et al.*, 2000). In addition, mice lacking p27<sup>Kip1</sup> have an extended number of progenitor proliferation cycles. These cells later differentiate leading to larger cell numbers in different parts of the body (Fero *et al.* 1996; Levine *et al.*, 2000).

As early as E11.0 (Drager, 1985; Young, 1985), a population of RGCs emerge to begin formation of the GCL in the mature retina. RGC axons fasciculate and extend their axons towards the optic stalk, forming the optic nerve, and first emerge at the optic chiasm at E12.5 (Silver, 1984). At the optic chiasm, the RGC growth cones encounter a midline. Here, about 20% of the growth cones in mice crossover to the contralateral side of the optic chiasm, while the remaining population extend along the ipsilateral side before innervating neurons in the LGN and superior colliculus (Drager and Olsen, 1980). In stereoscopic visual systems, such as those found in primates, the number of RGCs which cross the midline is about 50% (McIlwain, 1996).

Two homeodomain proteins, Brn-3b and Pax2, are important for normal RGC and optic chiasm development, respectively. Brn-3b is essential for the development of a large subset of RGCs. Mice lacking Brn-3b lose about 70% of their RGCs compared to their wild type littermates (Gan *et al.*, 1996), and defects in axon bundling are present among the remaining RGCs (Gan *et al.*, 1999). In addition, mice lacking Pax2 have RGCs which fail to crossover at the optic chiasm (Alvarez-Bolado *et al.*, 1997).

The majority of the neuroretinal cells at E12.5, however, are progenitors which continue to proliferate, increasing the size of the neuroblastic layer. The molecules which regulate neuroretinal progenitor cell proliferation are generally unknown. One regulatory protein which has been identified, however, is the homeodomain protein Chx10. Chx10 function is required for retinal progenitor proliferation and is present highly in the

neuroblastic layer of the presumptive neuroretina and other parts of the CNS, beginning at about E9.5 (Liu *et al.* 1994). Loss of Chx10 function results, both in mouse (Burmeister *et al.*, 1996) and human (Ferda-Pencin *et al.*, 2000), in microphthalmia (small eye) and a hypocellular retina, a consequence of impaired retinal progenitor cell proliferation. In inbred mice (strain 129/Sv) Chx10 loss-of-function is associated with optic nerve aplasia, but optic nerves do form when the mutation is expressed in other strains (Bone-Larson *et al.*, 2000), and human *CHX10*<sup>-/-</sup> subjects (Ferda-Pencin *et al.*, 2000).

As the presumptive neuroretina continues to develop and proliferate, other morphologically distinct cells begin to appear. Beginning at ~E12.0, axonless horizontal cells and committed but undifferentiated cone photoreceptors become post-mitotic. Between E11.5-12.5, the first amacrine cells are born, followed by the births of rod photoreceptors (~E14.0) and bipolar (~E17.0) cells (Young, 1985). By about E19.0, the RGC, INL, and ONL become well defined, although rod photoreceptors continue to develop up until about post-natal day 10 (P10), and outersegments do not fully form until 4-6 weeks after birth (Carter-Dawson and LaVail, 1979; Obata and Usukura, 1992).

Little is known about the molecular regulation of horizontal and amacrine cell development. In the adult mouse retina, Chx10 is expressed throughout the INL, including amacrine cells, but most abundantly in bipolar cells (Liu *et al.*, 1994). In adult *Chx10*<sup>-/-</sup> mice, amacrine cells are present, but, bipolar cells are absent, indicating that Chx10 is required for bipolar cell specification, differentiation, or maintenance (Burmeister *et al.*, 1996).

In contrast, photoreceptor development is better understood. Several compounds which affect photoreceptor determination and differentiation *in vitro* have been identified,

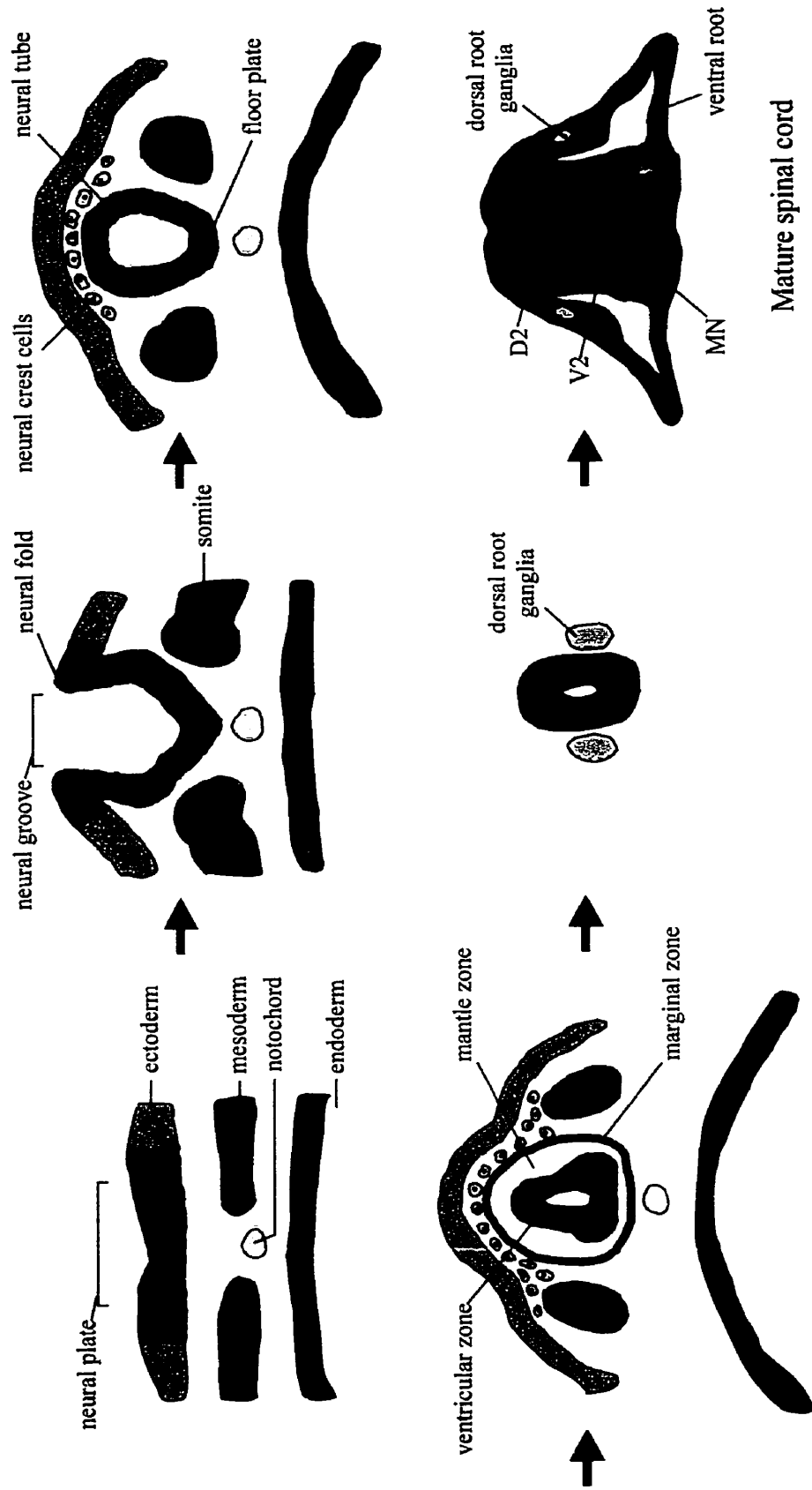
including retinoic acid (Kelley *et al.* 1994), S-laminin (Hunter *et al.* 1992), and Sonic hedgehog (Levine *et al.* 1997).

### **Mouse spinal cord development**

The spinal cord can be thought of as a biological information superhighway, establishing neural connections between the central nervous system (CNS) and the peripheral nervous system (PNS). Connected to the caudal CNS, the spinal cord consists of a long column of neurons running along the rostral-caudal axis with numerous lateral axon projections for PNS communication.

The spinal cord develops from the neural tube (composed of ectodermal tissue), by a process called neurulation. In mice, neurulation begins at about embryonic day 8.0-5 (E8.0-5) (Kaufman, 1992) near the anterior portion of the embryo, by the future midbrain, and proceeds along the anterior-posterior axis (Wolpert, 1998; Golden, 1993). There are two modes of neurulation (primary and secondary) which occur, depending on the location along the axis of the embryo. Primary neurulation occurs towards the anterior portion of the embryo. Here, ectoderm begins to fold to form a hollow column called the neural tube. Secondary neurulation occurs along the dorsal region of the embryo. In this process, a solid rod is formed initially. The rod develops into a neural tube after the generation of a lumen within the rod (Gilbert, 1997).

The development of the neural tube requires an elongation of the ectodermal midline cells. This morphological change causes a rising of the cells along the midline, resulting in a depression of the midline cells thereby creating the neural plate. The edges of the neural plate thicken and move towards each other to form the neural fold, while the



**Figure 1-3.** Formation of the vertebrate spinal cord (primary neurulation illustrated).

Blue regions in the mature spinal cord labeled D2, V2, and MN (motoneurons) indicated where their respective nuclei are located. (Figures adapted from Hall, 1992; Gilbert, 1997).

neural plate takes on a U-shape, becoming the neural groove. The neural folds continue to migrate towards the midline and eventually fuse, forming the neural tube. Once fusion of the neural folds occurs, the neural tube detaches from the ectodermal cells. These ectodermal cells will later become the epidermis (Schoenwolf and Smith, 1990; Wolpert, 1998).

The neuroblast cells that line the luminal surface of the neural tube make up the ventricular zone. The mitotic neuroblasts proliferate and differentiate into neurons and glial cells which migrate away from this zone (Hall, 1992). As the neuroblasts continue to proliferate and differentiate, the diameter of the lumen decreases and it is nearly obliterated in the mature spinal cord (Kaufman, 1992). Some of the molecular signals involved in neural patterning of the neural tube are believed to originate from the most ventral region of the tube, known as the floor plate, and from the underlying mesoderm, called the notochord. The notochord is a rod-shaped mesodermal structure running along the rostral-caudal axis, lying ventral to the neural tube (Gilbert, 1997).

Neural crest cells, originating from the tips of the neural folds, migrate into the dorsolateral regions of the neural tube. There, they differentiate into sensory ganglion cells, post-ganglionic autonomic neurons, peripheral glial cells, and epidermal pigment cells (Bronner-Fraser and Fraser, 1989; Gilbert, 1997).

Following neural tube closure, constriction occurs, forming the different chambers of the brain and the spinal cord. Progenitor cells differentiate into neurons and glial cells. Simultaneously, cellular migration and rearrangements occur, helping to define the different functional regions of the central nervous system. The most rostral portion of the neural tube will develop into the brain, forming the telencephalon, diencephalon,



mesencephalon, metencephalon, and myelencephalon, while the remaining part of the neural tube, caudal to the developing brain, will give rise to the spinal cord (Gilbert, 1997).

Once the spinal cord has formed, it serves as a communication link between sensory organs (*e.g.* through neurons found in the dorsal root ganglia), the brain, and motor network (*e.g.* through the motor neurons in the spinal cord that extend axonal processes through the ventral roots).

### **Dorsoventral patterning of the spinal cord**

In the mature spinal cord, different classes of neurons are regionalized symmetrically about the midline that runs along the dorsoventral axis of the spinal cord. For example, the nuclei of commissural neurons, which extend axons across the midline, are situated in the dorsal portion of the spinal cord (Gilbert, 1997), while nuclei of motoneurons are generally located in the ventral horns (Eisen, 1999). Sensory neurons, derived from neural crest cells, are located lateral to the dorsal spinal cord and form the dorsal root ganglia (DRG) (Gilbert, 1997).

During mouse spinal cord development, the first post-mitotic neurons are born at around E9.5 and neurogenesis continues until ~E14.5 (Nornes and Carry, 1978; Sims and Vaughn, 1979). Post-mitotic neurons migrate away from the ventricular zone and begin to project axons. For example, at E9.5 dorsal commissural interneurons, initially located lateral to the roof plate, begin to migrate ventrally away from the roof plate. Between E10.5 and E11.5, these neurons project axons toward the floor plate. These axons cross the floor plate midline at ~E12.5, and continue to extend into the ventral and lateral

funiculi (Zou *et al.*, 2000; Imondi *et al.*, 2000). Other neurons, such as tochlear motoneurons, differentiate and extend axons between E9.5 and E11.5 along a ventral-to-dorsal trajectory, away from the floor plate (Fritzsche *et al.*, 1995). Motoneurons in the ventral horns project axons into the periphery at around E11.5 (Chen *et al.*, 2000) where they will ultimately connect with muscle fibres. Sensory neurons in the DRG send axons to the dorsal spinal cord. Initially at E11.5, these axons navigate towards the dorsal root entry zone (DREZ) located dorsolateral to the roof plate cells. Upon arrival at the DREZ, the axons stall along the periphery of the dorsal spinal cord until about E14.5, when a subset of sensory axons (the proprioceptive afferents) enters the dorsal spinal cord and extends into the ventral horns. Other sensory axons, such as the noci-/thermoceptive afferents, remain at the DREZ, while mechanoceptive afferents migrate only into the dorsal horns (Pushcel *et al.*, 1996). Therefore, wiring of the spinal cord involves axon projections that must travel variable distances.

Cell patterning of the neural tube is believed to be controlled by ventral signals originating from the notochord and the floor plate cells in conjunction with signals from the dorsal epidermal ectoderm (Yamada *et al.*, 1993). *Sonic hedgehog* (*Shh*), expressed in the notochord and the floor plate, is a secreted ligand that forms a concentration gradient along the dorsoventral axis of the spinal cord (Eisen, 1999; Gilbert, 1997). *Shh* is important for the generation of ventral cell types such as motor neurons, since ectopic expression of *Shh* from a donor floor plate graft can induce the formation of ectopic motor neurons (Yamada *et al.*, 1993). In the dorsal neural tube, genes including *Pax3*, *dorsalin*, and *Msx1*, encode proteins with dorsalizing activity. *Shh* represses the expression of these genes in the ventral region of the neural tube (Liem *et al.*, 1995). In

the absence of Shh, the expression of *Pax3*, *dorsalin*, and *Msx1* expand ventrally and cells located towards the ventral neural tube take on a more dorsalized fate (Gilbert, 1997). *Bmp4* and *Bmp7*, members of the transforming growth factor beta (TGF $\beta$ ) superfamily of ligands, are inhibited by high levels of Shh. The two proteins are present in the presumptive dorsal epidermis (where levels of Shh are low), and are able to antagonize the effects of Shh by promoting the expression of *Pax3* and *Msx1* in the dorsal neural tube. Therefore, in the spinal cord, *Bmp4*, *Bmp7*, and Shh provide positional cues with opposing activity from opposite poles of the dorsoventral axis (Gilbert, 1997).

The mechanism by which progenitor cells interpret the extracellular gradient of Shh protein is not clearly understood. Three zinc finger transcription factors, *Gli1*, *Gli2*, and *Gli3*, are components of the hedgehog signaling pathway in vertebrates. The levels of *Gli* gene expression are affected by different concentrations of Shh (Ingham, 1998), and it has been proposed that the different levels of *Gli* may regulate the expression of downstream transcription factors which may specify different classes of neurons (Briscoe *et al.*, 2000). Surprisingly, targeted disruption of *Gli1* or *Gli3* does not drastically affect the neuronal patterning in the ventral spinal cord, while *Gli2* mutants fail to differentiate floor plate and V3 interneurons (Ding *et al.*, 1998; Matise *et al.*, 1998; Park *et al.*, 2000; Littingtung and Chiang, 2000). Therefore, it is possible, in the ventral spinal cord, additional transcription factors are involved in interpreting the Shh protein gradient (Briscoe *et al.*, 2000).

Recently, the neuronal fate of a subset of progenitor cells in the ventral neural tube has been examined (Briscoe *et al.*, 2000). The homeodomain proteins *Nkx6.1*, *Nkx2.2* (both induced by Shh) (Brisco *et al.*, 1999), and *Irx3* (repressed by Shh) (Briscoe

*et al.*, 2000), for example, can act in combination to specify different neuronal identities. Nkx2.2 can induce V3 neurons. Nkx6.1, in the absence of Irx3, leads to motor neuron development, while Nkx6.1, in the presence of Irx3, results in V2 neurons (Briscoe *et al.*, 2000). Other transcription factors which are controlled by Shh levels, including Pax6, Pax7 (Goulding *et al.*, 1993; Ericson *et al.*, 1997), Dbx1, and Dbx2, may be important in the development of other classes of neurons.

The molecular details of eye and spinal cord development are slowly being determined and through the concerted effort of the Human Genome Project, molecules important in the development of the nervous system are being discovered at an unprecedented rate (Boguski *et al.*, 1993; Boguski *et al.*, 1994). By determining the function of the genes involved in neural development, we will move closer towards understanding how the human brain develops and functions.

#### **dbEST: A source for novel genes**

The human genome is estimated to contain about 35,000 to 120,000 genes (Ewing and Green, 2000; Liang *et al.*, 2000) but a typical cell expresses only a subset of these genes. For example, neurons express ~25,000 genes, while epithelial cells express ~15,000 genes (Velculescu *et al.* 1999). As a first step towards identifying the subsets of genes which are expressed in different tissues *en mass*, cDNA libraries from different tissues were generated, and partial, single-pass cDNA sequences were obtained from randomly picked clones as a part of the Human Genome Sequencing project (Boguski *et al.*, 1993). To reduce the redundancy of highly expressed genes, the majority of these libraries have been normalized to maximize the likelihood of identifying unique gene sequences (Bonaldo *et al.*, 1996). These sequences (ESTs) have been a major source for

novel gene discovery and have been deposited into a database called dbEST, which contains ESTs from many different organisms (Boguski *et al.*, 1993). Currently, human ESTs contribute the majority of dbEST entries (with 2,275,000), although there are also a number of ESTs from model organisms including *Mus musculus* (1,632,000), *Rattus* sp. (188,000), and *C. elegans* (101,000) (dbEST, <http://www.ncbi.nlm.nih.gov/dbEST>, Sept. 2000). Since EST entries include information about their library of origin, one useful resource dbEST provides, in conjunction with *basic local alignment search tool* (BLAST, <http://www.ncbi.nlm.nih.gov/BLAST>) (Altschul *et al.*, 1990), is the ability to examine the approximate tissue expression profile of a sequence of interest (more commonly referred to as a “virtual northern”). Many ESTs have also been clustered together based on their overlapping sequences to generate “virtual” expressed sequences (UniGene, <http://www.ncbi.nlm.nih.gov/UniGene>).

UniGene is a database that can display clusters of ESTs originating from only one tissue source (Wheeler *et al.*, 2000). These clusters are therefore enriched for novel, tissue-specific genes which are likely to be important for the function of that particular tissue. Although the majority of the EST clusters will be 3’UTR biased, some clusters may have a partial open reading frame (ORF) encoding a domain that can provide a clue towards the function of the particular gene.

## The CUB domain

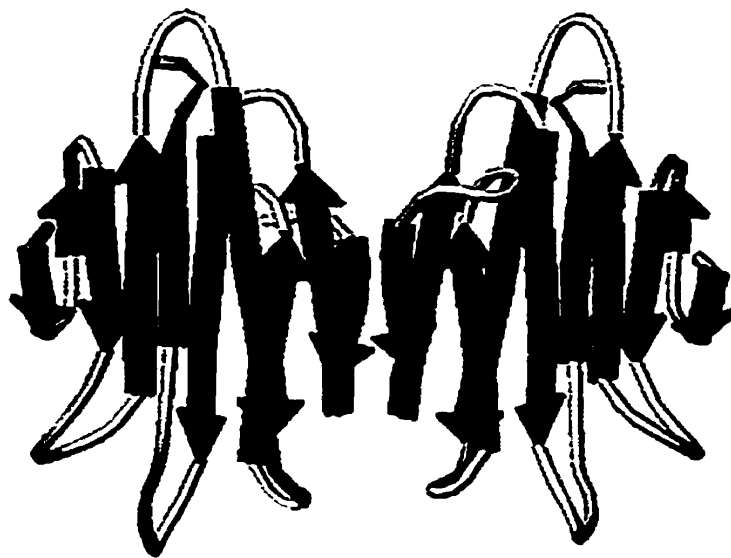
The structure and function of proteins are the consequence of their amino acid sequences. In some proteins, these amino acid sequences are conserved both in function and evolution. For example, the homeodomain, which is highly conserved between invertebrates and vertebrates, encodes a helix-turn-helix motif found in many transcription factors. This motif functions to bind double stranded DNA (Otting *et al.*, 1990). Some conserved motifs mediate protein-protein association. For example, leucine zipper motifs, which form alpha helical structures containing periodical leucine residues, often function in protein-protein dimerization (Moran *et al.*, 1994). By shuffling these modular domains in different combinations, proteins with different biochemical properties can rapidly evolve (*e.g.* through alternative splicing, gene duplication and/or recombination) from a limited initial reservoir of protein sequences. Although this is very much simplified, the result of these combinations can generate molecules which may be uniquely adapted to specific biochemical and cellular processes. Homeodomain proteins, for example, are transcription factors which play critical roles in development (Duboule, 1994). Another domain found in proteins often involved in developmental processes is the CUB domain (Bork and Beckmann, 1993).

The CUB domain, originally identified in complement subcomponents C1r/C1s (Leytus *et al.*, 1986; Tosi *et al.*, 1987), *Uegf* (Delgadillo-Rosso *et al.*, 1989), and *Bmp1* (*bone morphogenetic protein 1*)(Wozney *et al.*, 1988), is an extracellular domain ranging from 110 to 133 amino acids. It is involved in protein-protein (Bork and Beckmann, 1993) or protein-carbohydrate interactions (Solis *et al.* 1998). CUB domain-containing proteins are functionally diverse, found in both invertebrates and vertebrates, and are

often involved in developmentally regulated processes (Bork and Beckmann, 1993). Currently, there are over 36 CUB domain proteins (Prosite, <http://expasy.cbr.nrc.ca>, Sept. 2000) some of which are listed in Table 1-1. Proteins containing CUB domains often have other domains such as a metalloprotease domain and/or epidermal growth factor (EGF) - like repeats. CUB domains, however, can also function alone as illustrated by the spermadhesins, a family of genes that encode a protein consisting of a signal sequence and a single CUB domain. Spermadhesins are secreted into the extracellular matrix and play a critical role in fertilization (Solis *et al.*, 1998) by mediating the attachment of the sperm cell to the zona pellucida.

The primary sequence of CUB domains is not highly conserved. CUB domains of different proteins generally share only about 25% amino acid identity (Bork and Beckmann, 1993), although the conservation between CUB domains within the same protein family can be as high as 40%.

The most conserved residues in CUB domains are four cysteine residues (C1-C4) which are important for the formation of two disulfide bridges (between C1-2 and C3-4). The crystal structure of several spermadhesin dimers, including aSFP (Dias *et al.*, 1997) (Figure 1-4) and PSP-1/PSP-II (Romero *et al.*, 1997), has been solved. CUB domains consist of two  $\beta$ -sheets, each composed of five  $\beta$ -strands. The  $\beta$ -turns connecting the  $\beta$ -strands are the most variable portion of CUB domains and are exposed at the surface where they are available for interaction with other molecules. The overall topology of the CUB domain resembles the antigen binding region of immunoglobulins (Dias *et al.* 1997; Bork and Beckman, 1993) or a “jellyroll” (Romero *et al.*, 1997).



**Figure 1-4.** Crystallographic dimer of two aSFP spermadhesin molecules

Each aSFP spermadhesin monomer is composed of a single CUB domain. The CUB domain consists of 10  $\beta$  strands (blue arrows) forming two  $\beta$  sheets. Spermadhesin dimerization is stabilized by hydrogen bonds between  $\beta 8$  and  $\beta 7'$ , and  $\beta 7$  and  $\beta 8'$  strands (the ' denotes the second monomer). In other spermadhesin dimers, the interactions may occur between  $\beta 7$  and  $\beta 7'$ , and  $\beta 8$  and  $\beta 8'$  (e.g. between PSP-I/PSP-II heterodimers). (Figure from Romero *et al.*, 1997).



**Table 1.1.** Summary of various CUB domain-containing proteins

Protein	Organism	Domains	Notes
hensin	mammals	8 SRCR (scavenger receptor, cysteine rich) 2 CUB, 1 ZP (zona pellucida)	possible role in determining polarized phenotype of some epithelia and brain cells (Takito <i>et al.</i> , 1999)
CRP-ductin/ vomero glandin	mouse	8 SRCR, 3 CUB, 1 ZP	unknown role in pheromone perception (Matsushita <i>et al.</i> , 2000)
ebnerin	rat	4 SRCR, 3 CUB, 1 ZP	identified in a screen for novel taste bud receptor genes (Asano-Miyoshi <i>et al.</i> , 1998)
DMBT1	human	9 SRCR, 2 CUB, 1 ZP	frequently deleted at 10q25-26 malignant gliomas, alt-splice form of hensin (Takito <i>et al.</i> , 1999)
hch-1	<i>C. elegans</i>	1 astacin metalloprotease domain, 1 CUB, 1 EGF	required for normal hatching and post-embryonic cell migration (Hishida <i>et al.</i> , 1996)
C1s/C1r	mammals	2 CUB, 2 Coagulation factor-like, 1 EGF, trypsin serine protease domain	C1s, C1r, and C1q form C1 complex in complement cascade (Thielens <i>et al.</i> , 1999)
enteropeptidase	vertebrate	2 low-density lipoprotein receptor-like domains, 1 MAM, 1 SRCR, 1 CUB	intestinal membrane protein needed for activation of trypsinogen (Yuan <i>et al.</i> , 1998)
Bmp1	vertebrate	1 astacin metalloprotease domain, 3 CUB, 2 EGF	process procollagen; cleaves chordin (Scott <i>et al.</i> , 1999); ectopic bone formation (Wozney <i>et al.</i> , 1988)
Tolloid Tolkin	drosophila	1 astacin metalloprotease domain, 5 CUB, 3 EGF	cleaves sog; dorsal ventral patterning (Marques <i>et al.</i> , 1997) embryonic lethal (Finelli <i>et al.</i> , 1995)
mammalian tolloid	mammals	1 astacin metalloprotease domain, 5 CUB, 3 EGF	process procollagen (Scott <i>et al.</i> , 1999)
tolloid-like		as above	process procollagen, cleaves chordin (Scott <i>et al.</i> , 1999)
tolloid-like2		as above	unknown (Scott <i>et al.</i> , 1999)

**Table 1.1.** (continued)

Protein	Organism	Domains	Notes
BP10/SpAN	sea urchin	1 zinc-metalloprotease domain, 1 EGF, 2 CUB	secreted at blastula stage by sea urchin embryo (Lhomond <i>et al.</i> , 1996)
F42A10.8 R151.5	<i>C. elegans</i>	1 zinc-metalloprotease domain, 1 EGF, 1 CUB	hypothetical ORF (NCBI) hypothetical ORF (NCBI)
K03E5.1	<i>C. elegans</i>	2 CUB	hypothetical ORF (NCBI)
neuropilin-1/2	vertebrates	2 CUB, 2 Coagulation factor-like, 1 MAM, 1 TM	axon guidance, receptor for type III semaphorins; receptor for VEGF isoform (Soker <i>et al.</i> , 1998)
Fibropellins I and III (Uegf)	sea urchin	1 CUB, variable no. of EGF, C-ter avidin-like domain	form fibers that surround the surface of the embryo (Bisgrove <i>et al.</i> , 1991)
TSG-6/ PS4	mammals	1 link domain, 1 CUB	a serum and growth factor which binds to hyaluronate (Day, A.J. 1999)
spermadhesins	mammals	1 CUB	bind carbohydrates; coat sperm cell, allowing for adhesion to egg (Romero <i>et al.</i> , 1997)
UVS.2/ hatching enzyme (HE)	Xenopus	1 astacin metalloprotease domain, 2 CUB	expressed during dorsoanterior development (Katagiri <i>et al.</i> , 1997); egg hatching
AEA	cray fish	1 astacin metalloprotease domain, 2 CUB	expressed in unhatched juveniles prior to hatching; no expression in adult (Geier and Zwillig, 1998); egg hatching
Cubulin	mammals	27 CUB, 8 EGF N-ter membrane association segment (non-TM)	vitamin B <sub>12</sub> receptor; HDL receptor (Kristiansen <i>et al.</i> , 1999)

## **Tolloid and tolloid-related proteins**

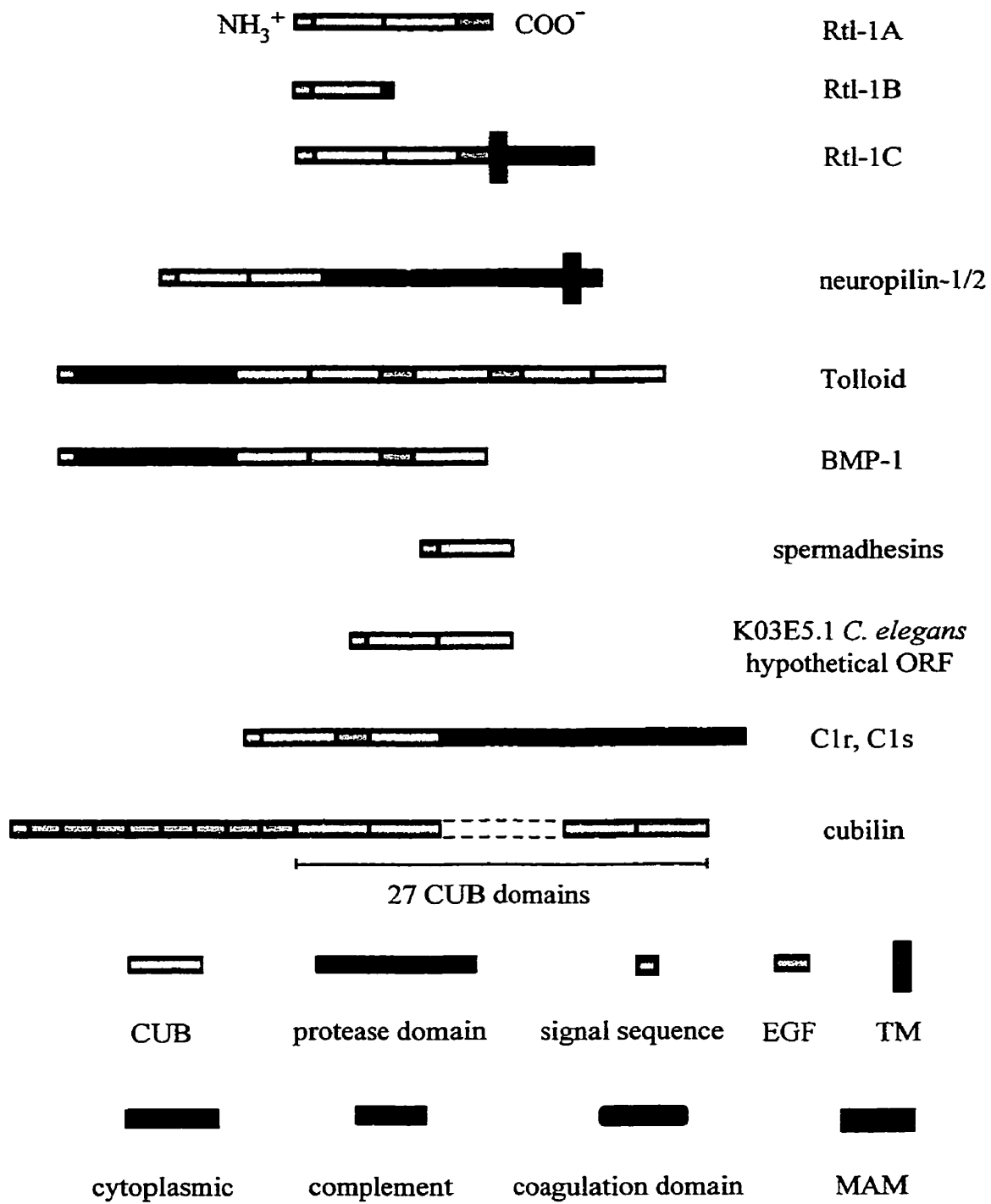
A class of CUB domain-containing proteins to which *RTL-1* shares resemblance with is the tolloid family of proteins, whose prototype members are tolloid (Shimell *et al.*, 1991) and Bmp1 (Wozney *et al.*, 1988). Tolloid and Bmp1 are secreted proteins with an astacin-like metalloprotease domain, multiple CUB domains, and EGF-like repeats. Tolloid (*tld*) is a *Drosophila melanogaster* protein important for dorsoventral patterning of the embryo. Mutations in *tld* result in a partial transformation of the dorsal ectoderm into ventral ectoderm (Shimell *et al.*, 1991). This phenotype is similar to null mutants of *decapentaplegic* (*dpp*, a *Drosophila* protein related to Bmp2 and Bmp4) suggesting that *tld* and *dpp* act in dorsoventral patterning (Shimell *et al.*, 1991).

Biochemical studies on a secreted *Drosophila* protein, short gastrulation (*sog*), put *dpp* and *tld* together in the same pathway. *Sog* binds to *dpp*, creating inactive *dpp* complexes, thereby antagonizing *dpp* signaling (Schmidt *et al.*, 1995). *Tld* can bind to *dpp-sog* complexes and cleave *sog*, which releases *dpp* (Marques *et al.*, 1997). Thus, one function of *tld* is to regulate the antagonistic effects of *sog* on *dpp*.

This biochemical pathway is also conserved in vertebrates. In *Xenopus*, the protein Xolloid (*Xld*), orthologous to tolloid, can bind and cleave Chordin (*Chd*), a protein orthologous to *Sog*. Proteolytic cleavage of *Chd* prevents it from binding to Bmp4 (homologous to *dpp*), which allows the free Bmp4 to bind and activate its receptor (Piccolo *et al.*, 1997).

In mammals, three proteins with a domain organization identical to Tolloid have been identified: mammalian Tolloid (*mTld*) (Takahara *et al.*, 1994), mammalian Tolloid-like-1 (*mTll-1*; previously known as *mTll*) (Takahara *et al.*, 1996) and mammalian

Tolloid-like-2 (mTll-2) (Scott *et al.*, 1999). A shorter protein, Bmp-1, is derived as an alternative splice product from *mTld* (Takahara *et al.*, 1994). The mammalian tolloids are expressed widely among different tissues in both the developing and mature mouse (Takahara *et al.*, 1994; Takahara *et al.*, 1996; Scott *et al.*, 1999). Targeted disruption of murine *mTld/Bmp1*, although embryonic lethal, does not lead to severe malformations in axial patterning (Suzuki *et al.*, 1996) as observed in *Drosophila tld* mutants (Shimell *et al.*, 1991). These mouse embryos die from a herniated gut (Suzuki *et al.*, 1996), and the lack of any dorsoventral axial defects suggested other mammalian tolloids may have functional redundancies (Scott *et al.*, 1999). Mice in which *mTll-1* is disrupted, like the *mTld/BMP1* mouse mutant, also lack axial defects. Instead, *mTll-1*<sup>-/-</sup> mice are embryonic lethal due to defects confined to the heart (Clark *et al.*, 1999). A mouse knockout of *mTll-2* has not yet been reported. Recent biochemical analysis of these metalloproteases have provided insight into the possible roles of these proteins in mouse development. Bmp1, mTld, and mTll-1, but not mTll-2, are capable of processing procollagen-C propeptides. In addition, Bmp1 and mTll-1 are capable of cleaving chordin, and this biochemical redundancy may explain the lack of axial defects when either *mTld/Bmp1* or *mTll-1* are disrupted, as both genes have regions of overlapping expression patterns (Scott *et al.* 1999). Despite the lack of axial defects, these proteins may play a major role in antagonizing chordin activity.



**Figure 1-5.** Organization of various CUB domain-containing proteins

## Neuropilins

The other class of CUB-containing proteins to which *RTL-1* resembles is the neuropilins. A small family of type I transmembrane proteins, neuropilins are capable of binding to class 3 semaphorins (He and Tessier-Lavigne, 1997; Kolodkin *et al.*, 1997). Semaphorins, a family of secreted and transmembrane proteins, are involved in axonal guidance and serve as attractive (Bagnard *et al.*, 1998) or repulsive cues (Puschel, 1996). Currently, there are two members of the neuropilin family (neuropilin-1 and neuropilin-2), that share about 44% amino acid identity with each other. The domain organization of the two neuropilins is identical. Both have two CUB domains, two coagulation factor-like domains, a MAM domain, a transmembrane (TM) region, and a short cytoplasmic tail. In binding assays, the CUB domains of neuropilins were shown to be the major site for semaphorin binding and specificity (*e.g.* swapping the CUB domains of neuropilin-1 and neuropilin-2 is sufficient to change the semaphorin response of the respective receptors) (Nakamura *et al.*, 1998). The coagulation factor-like domain is homologous to a ~160 amino acid region of coagulation factor V and VIII. In neuropilin-1, this domain provides a binding site for the basic carboxy-terminus of Sema3A (Nakamura *et al.*, 1998). The MAM domain was originally identified as a 220 - 240 amino acid motif in the proteins meprin, A5, and  $\mu$ , and has been shown to be required for neuropilin-1 receptor multimerization and signal transduction (Nakamura *et al.*, 1998; Giger *et al.*, 1998).

Neuropilin-1 is expressed in different regions of the developing and mature nervous system, including parts of the brain, RGCs within the eye, and some non-neuronal tissues. In the peripheral nervous system (PNS), neuropilin-1 is expressed in the

trigeminal, facial, glossopharyngeal, and vagus nerves, and in the spinal sensory and motoneurons (Kawakami *et al.*, 1996). Expression of neuropilin-1 has been localized to axons and growth cones (Fijusawa *et al.*, 1995; Takagi *et al.*, 1995; Kawakami *et al.*, 1996), and ectopic over-expression of neuropilin-1 in the mouse embryo causes defasciculation and altered sprouting of axon fibers (Kitsukawa *et al.*, 1995).

Neuropilin-2 is also expressed in diverse parts of the nervous systems, including areas overlapping with neuropilin-1, such as motoneurons and RGCs, and in distinct regions including cerebellum and sympathetic neurons. In addition, neuropilin-1 is highly expressed in the dorsal root ganglia (DRG), where neuropilin-2 expression is almost undetectable (Chen *et al.*, 1997).

Neuropilin-1 and -2 can form both non-covalent homodimers, and in cells where they are coexpressed, they have been shown to form heterodimers, which is believed to confer different ligand binding specificities (Takahashi *et al.*, 1998). Ligand binding experiments show that neuropilin-1 has high binding affinity for Sema3A, Sema3D, and Sema3E (He and Tessier-Lavigne, 1997), while neuropilin-2 has high binding affinity for Sema3C and Sema3F, but low binding affinity for Sema3A (Chen *et al.*, 1997; Giger *et al.*, 1998). Neuropilins have short cytoplasmic regions (about 40 amino acids) and alone, cannot propagate a semaphorin signal into the cell. Therefore, to transduce the semaphorin signal into the cell, Kolodkin and Ginty (1997) have suggested that a functional semaphorin receptor includes neuropilin as part of a larger protein complex.

Recently, it has been shown that *Drosophila* plexin A, a member of a large family of transmembrane proteins found in organisms ranging from viruses to humans, is a receptor for Sema-1a and Sema-1b (Winberg *et al.*, 1998, Tamagnone *et al.*, 1999). In

mice, plexin-1 (a homolog of plexin A) cannot directly bind Sema3A, but is able to complex with neuropilin-1 to form a functional Sema3A receptor complex (Takahashi *et al.*, 1999). Formation of a neuropilin-1/plexin-1 complex enhances the binding affinity of neuropilin-1 for Sema3A as compared with neuropilin-1 alone. Dominant negative mutants of plexin-1 which lack the intracellular tail can still bind to neuropilin-1, but cannot transduce the Sema3A signal in DRG neurons. This finding suggests that in a plexin-1/neuropilin-1 receptor complex, neuropilin-1 serves to bind Sema3A while plexin-1, which has a longer cytoplasmic region than the neuropilins, functions to transduce the Sema3A signal into the cell (Takahashi *et al.*, 1999).

Neuropilin-1 has also been linked to blood vessel morphogenesis. It has been recently shown that neuropilin-1 is also a component of the VEGF165 receptor, an isoform of vascular endothelial growth factor (VEGF). VEGF is important for the regulation of endothelial cell proliferation, vasculogenesis, angiogenesis, and vascular permeability (Soker *et al.*, 1998, Kawasaki *et al.*, 1999).

Neuropilin-1 deficient mice are embryonic lethal, dying between E10.5 and E12.5 from cardiovascular defects, which may be due to disruption of VEGF165 signaling. These mice also have anomalies in axon efferent trajectories of cranial, spinal, and trunk level nerves which are generally defasciculated (Kitsukawa *et al.*, 1997).

Unlike neuropilin-1-deficient mice, neuropilin-2 deficient mice are viable and live into adulthood. In the absence of neuropilin-2, there is disruption in the fasciculation and organization of different cranial and spinal nerves (Giger *et al.*, 2000). In addition, nerve tracts including the mossy fibers of the hippocampus, and anterior commissures have axon migration defects (Chen *et al.*, 2000).



## **CHAPTER 2**

Cloning and characterization of *Rtl-1*:  
A novel CUB-encoding gene expressed in  
the developing and mature mouse nervous system

## INTRODUCTION

The molecular components required for mammalian eye development are gradually being identified as novel genes are discovered. Through the Human Genome Project, thousands of new genes, represented as ESTs, have been found. Despite this plethora of data, the identity of the vast majority of these genes is unknown.

The EST databases represent a formidable source for gene discovery. As a simple *in silico* screen to identify novel retinal genes potentially involved in eye development, I scanned the UniGene database for retina-only EST clusters with sequence homology to proteins involved in eye development. One of these clusters, with homology to the CUB domains found in tolloids and neuropilins, was named *RTL-1*.

A human *RTL-1* cDNA was isolated by screening a high density filter array of a normalized human retinal cDNA library using a 3'UTR DNA probe that was PCR-amplified from human genomic DNA. This human cDNA, encoding a partial ORF, was subsequently used to screen three different mouse cDNA libraries to obtain a full-length murine ortholog of *RTL-1*. The murine *Rtl-1* cDNA sequence was used to query the sequence databases and three human *bacterial artificial chromosomes* (BACs) were identified that were used to reconstruct a full-length *RTL-1* ORF.

To determine whether *RTL-1* mapped to any human disease loci, radiation hybrid (RH) mapping was performed in collaboration with Dr. S.W. Scherer, and the corresponding syntenic regions were examined using mapping data available at The Jackson Laboratories (<http://www.jax.org>). As a first step towards understanding the biological role of *Rtl-1*, the expression pattern in adult and developing mouse was determined using multi-tissue northern blots and *in situ* hybridization, respectively.

## **MATERIALS AND METHODS**

### ***In silico* screen**

A subset of retina-only clusters from UniGene libraries 177 (N2b5HR) and 178 (N2b4HR) represented by >1 EST was compared with the non-redundant (nr) database using BLASTN and BLASTX, both variants of BLAST, (Altschul *et al.*, 1990) to identify conserved sequences. UniGene clusters with homology to known protein domains or motifs, particularly those which are often found in proteins important in development, were further investigated. dbEST was also searched to obtain additional 5' sequence derived from the same clones ([http://www.ncbi.nlm.nih.gov/irx/dbST/dbest\\_query.html](http://www.ncbi.nlm.nih.gov/irx/dbST/dbest_query.html)).

### **PCR conditions**

PCR reactions were performed as described by Sambrook *et al.* (1989). Briefly, each reaction contained 400 ng of each primer, 200  $\mu$ M of each dNTP, 10 mM Tris-HCl pH 8.0, 50 mM KCl, 0.01% gelatin, 1.5 mM MgCl<sub>2</sub>, 8  $\mu$ g BSA, and ~1-10 ng of DNA template in a volume of 50  $\mu$ l overlaid with mineral oil. Samples were heated to 98°C for 10 min. and cooled to 80°C prior to the addition of 1  $\mu$ l (~10 u) *Taq* DNA polymerase. 30 thermal cycles consisting of a short denaturation at 94°C for 45 sec, 55 - 62°C annealing (depending on the primer pair) for 45 sec, and 72°C extension for 30 sec to 3 min. This step was followed by a 72°C final extension for 10 min and a 4°C incubation. PCR products were purified with QiaQuick PCR spin columns (Qiagen), and resuspended in 35 - 50  $\mu$ l of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

For subcloning purposes, *Pfu* DNA polymerase (Stratagene) was used. The reactions conditions were 400 ng of each primer, 200  $\mu$ M of each dNTP, 5  $\mu$ l of 10x cloned *Pfu* DNA polymerase reaction buffer (200 mM Tris-HCl pH 8.8, 20 mM MgSO<sub>4</sub>, 100 mM KCl, 100 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1% Triton X-100, and 1000  $\mu$ g/ml nuclease-free BSA) in a 50  $\mu$ l volume covered with mineral oil. The thermocycling program was identical to that described above except for a longer thermal cycle extension time of 6 to 10 min.

### **Radiolabeling**

Approximately 50 - 100 ng of DNA, gel purified using either QiaEx II beads or QiaQuick columns (Qiagen), was primed with random hexamers and labeled using Klenow and  $\alpha$ <sup>32</sup>P dCTP (Feinburg and Vogelstein, 1983). Labeling reactions were purified using Sepharose G50-150 spin columns, eluted in 100  $\mu$ l of TE, and incorporation of the radionucleotide was quantitated with a Beckman LS 6500 scintillation counter.

### **Southern blot transfers**

Southern blot transfers were performed essentially as described by Sambrook *et al.* (1989). DNA samples were resolved by electrophoresis using 0.5-0.8 % agarose gels containing 50 ng/ml ethidium bromide. Southern blot transfers involving DNA fragments larger than 10 kb were treated briefly with 0.2N HCl for 10 min and rinsed twice with ddH<sub>2</sub>O. The DNA was alkali denatured, neutralized, and transferred in 10x SSC by capillary action overnight onto Hybond-N+ (Amersham Life Science) nylon membrane

filters. The filters were cross-linked by UV irradiation, rinsed in 2x SSC, and air dried at room temperature.

### **Library screening**

To obtain a human *RTL-1* cDNA, 40,000 cDNAs from a normalized human retinal plasmid cDNA library (N2b4HR, previously transformed into *E.coli* DH5 $\alpha$  cells) were placed into 96 well microtitre plates containing LB broth with 100  $\mu$ g/ml ampicillin and 15% w/v glycerol. A Biomek 1000 (Beckman Instruments Inc.) robot was used to produce 25 high density nylon filter (Hybond-N, Amersham Life Science) arrays representing the 40,000 cDNAs. The filters, overlaying LB containing 1% agarose supplemented with 100  $\mu$ g/ml of ampicillin, were incubated overnight at 37°C. The filters were treated according to the manufacturer's directions, with 10% SDS, denatured in 0.5M NaOH, 1.5 M NaCl, and neutralized twice with 1M Tris-HCl pH 7.5. Following UV cross-linking, cell debris was removed by scrubbing the filters with KimWipes in 2x SSC, 0.1% SDS and rinsed in 2x SSC. The filters were hybridized in 50% formamide containing an  $\alpha^{32}$ P labeled probe ( $1 \times 10^5$  cpm/ml), representing 250 bp of the *RTL-1* 3'UTR that was PCR-amplified from human genomic DNA using primers RTL-F (5' CTCCTTCATTAAATCTATTATCT 3') and RTL-R (5' TCATATATGCATATTAGTGTACTG 3').

To identify an orthologous *Rtl-1* murine cDNA,  $1.8 \times 10^6$  plaque forming units (pfu) from a mouse developing eye (P1-P3) (courtesy of Dr. Jeremy Nathans, Howard Hughes Medical Institute, Baltimore, Maryland), adult mouse retina (a gift from Dr. Anand Swaroop, University of Michigan, Ann Arbor, Michigan), and adult mouse brain

(from Dr. Michael Hayden, University of British Columbia, Vancouver, British Columbia) phage cDNA libraries were screened with a PCR product generated using primers RTL-CDSF (5' CTTCAGAGATGCTGGCAGAG 3') and RTL-CDSR (5' CCATAGATTCCAGCTCTCCAT 3') encoding CUB1 from human *RTL-1*. For each 132 mm LB plate, approximately  $1 \times 10^5$  pfu was used to infect 400  $\mu$ l of *E. coli* LE392 or XL1-Blue cells grown exponentially to an OD<sub>600</sub> of 0.5 in LB supplemented with 0.4% maltose and overlaid using 1% LB top agarose. The plates were incubated overnight at 37°C, stored at 4°C for at least 2 hours, and duplicate filters for each of the 18 plates were prepared using Hybond N+ nylon membranes. The membranes were denatured in 0.5 N NaOH, 1.5M NaCl, neutralized in 1M Tris-HCl pH 7.5, 1.5 M NaCl, and rinsed in 2x SSC. The filters were UV cross-linked and then probed with a DNA fragment encoding CUB1 from human *RTL-1*. DNA hybridization and washing conditions were essentially as described below in the Hybridizations section, except that both prehybridization and hybridization solutions contained 40%, instead of 50%, formamide.

Agar plugs surrounding the positive plaques were picked and incubated in 500  $\mu$ l of SM buffer (0.1 M NaCl, 0.01 M MgSO<sub>4</sub>, 0.05 M Tris-HCl pH 7.5, and 0.01% gelatin) at 4°C overnight. Secondary, tertiary, and quaternary screens were performed using the same CUB1 DNA probe on filters lifted from serial dilutions of phage that produced ~100 pfu/plate. A total of 12 pure plaques were subsequently amplified and eluted in 200  $\mu$ l of SM buffer, with the addition of chloroform, and stored at 4°C.

## **Hybridizations**

UV cross-linked filters were briefly rinsed with 2x SSC, and prehybridized for at least 3 hours (for Southern blot hybridizations) or overnight (for northern blot hybridizations) at 42°C in 10 ml of prehybridization solution containing 40% or 50% formamide (for inter-species and intra-species DNA probes, respectively), 5x SSC, 1x Denhardt's solution, 1% SDS, 20 mM sodium phosphate buffer (pH 7.0), and 1 µg/ml sheared salmon sperm DNA. The hybridization solution was identical to the prehybridization mixture except for the addition of dextran sulphate to a final concentration of 10% w/v. Approximately  $5 \times 10^5$  to  $1 \times 10^6$  cpm/ml of denatured probe was added to the hybridization mixture and incubated for ~16 hrs at 42°C with gentle rotation.

## **Nylon membrane washing conditions**

Hybridized filters were washed once each in 2x SSC and 0.2x SSC for 10 min at room temperature (RT) with agitation, followed by a wash in 0.2x SSC, 0.1% SDS at 65°C for 1 hour, rinsed in 2x SSC at RT for 1 min, and exposed to X-OMAT-AR or BioMax MS (Eastman Kodak Company) film with the appropriate intensifier screens at -80°C. Additional washes were conducted as necessary.

## **Phage DNA isolation**

Phage DNA was isolated as described by Grossberger (1987). Briefly, *E. coli* LE392 cells were grown exponentially at 37°C to an O.D.<sub>600</sub> of 0.5. Approximately  $1 \times 10^6$  pfu of phage was used to infect 200 µl of LE392 cells mixed with 200 µl of

adsorption buffer (10 mM MgCl<sub>2</sub>, 10 mM CaCl<sub>2</sub>) and incubated at 37°C for 10 min. The cells were grown overnight with agitation at 37°C in 5 ml of LB supplemented with 0.2% glucose and 20 mM MgSO<sub>4</sub>. The cell debris was pelleted by centrifugation at 3,000 rpm at 4°C, and the supernatant was transferred into Ultra-Clear™ 14x89 mm tubes (Beckman Instruments Inc.). The phage was pelleted at 30,000 rpm for 30 min at 4°C using a SW41 rotor (Beckman Instruments Inc.), and resuspended in 400 µl of SM with 1 mg/ml proteinase K. After a 3 hour digestion at 37°C, 100 µl of 7.5 M ammonium acetate was added followed by a phenol/chloroform extraction and precipitated using 1 ml of 100% ethanol after incubation on dry ice for 20 min. The DNA was pelleted at 4°C by centrifugation at 12,000 rpm, rinsed with 1 ml of 70% ethanol, air dried, and resuspended in 30 µl of TE containing 200 µg/ml of RNaseA.

### **Isolation of total RNA**

Total RNA was isolated from adult C57BL/6 and embryonic mouse tissues. All mice were euthanized by cervical dislocation and tissues were dissected, rinsed in ice-cold 1x PBS, and homogenized in TRIzol Reagent (Gibco Life Technologies) as recommended by the manufacturer. The RNA was resuspended in diethylpyrocarbonate-treated distilled, deionized water (DEPC ddH<sub>2</sub>O) and quantified using a Beckman DU 530 spectrophotometer. RNA for use in RT-PCR was further treated with DNaseI. Approximately 10 µg of total RNA was resuspended in 16 µl of 1x first-strand DNA synthesis buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl<sub>2</sub>). DEPC ddH<sub>2</sub>O was added to a volume of 80 µl and incubated at RT with 1 µl (231 U/µl) of DNaseI (Gibco Life Technologies) for 15 min. The reaction volume was then raised to 100 µl



with the addition of DEPC ddH<sub>2</sub>O and treated with TRIzol. The purified RNA was resuspended in 20 µl of DEPC ddH<sub>2</sub>O.

### **Northern blot transfers**

Northern blots were prepared as essentially described by Sambrook *et. al.* (1989). All solutions were prepared with DEPC ddH<sub>2</sub>O. Between 5 - 10 µg of total RNA was resolved overnight at 43 volts (V) with an ethidium bromide stained 1% denaturing agarose gel (1% agarose, 1x MAE, 16.8% v/v formaldehyde, and 50 ng/ml ethidium bromide) in 1x MAE (40 mM MOPs, 10 mM sodium acetate, 1 mM EDTA, pH 7.0) running buffer. The gel was washed three times each for two hours with DEPC ddH<sub>2</sub>O, and transferred overnight by capillary action onto a Hybond N+ nylon membrane with 10x SSC. The membrane was UV cross-linked, rinsed with 2x SSC, and air-dried at room temperature.

### **First strand cDNA synthesis**

First strand cDNA was synthesized from mouse retina and whole mouse embryo by mixing 5 µg of total RNA (previously treated with DNaseI) with 0.5 µg of oligo d(T)<sub>12-18</sub> (Pharmacia Inc.) raised to a total volume of 12 µl with DEPC ddH<sub>2</sub>O. The mixture was incubated at 70°C for 10 min and cooled on ice for at least 2 min. A 7 µl cocktail containing 1x first strand DNA synthesis buffer, 28.5 mM DTT, and 1.4 mM dNTPs was added, mixed, and incubated at 42°C for 5 min. Afterwards, 1 µl of Superscript II RT (Gibco Life Technologies) (200 U) was added and the tube was incubated for an additional 50 min at 42°C. Following the reverse transcriptase reaction,

the enzyme was heat inactivated at 70°C for 15 min and cooled to 4°C. 1 µl of RNaseH (3.8 U) (Gibco Life Technologies) was added to the tube and incubated for 20 min at 37°C. For subsequent PCR reactions, 1 µl of the cDNA was used as template.

### **Subcloning**

Clone *RTLCD5*, encoding CUB1, was generated using an inverse PCR strategy. Primers H-INT3 (5' AAGCTTATTCCTTTAGTGAGGGTTAAT 3') and INRTL-F2 (5' TCAGGTGTGAAATTGTATCGAG 3') were used to amplify a ~3.5 kb fragment from clone *28RI63H3* (human) as a template with *Pfu* DNA polymerase. Clone *mRTL3UTR* was constructed using primers H-INT3 and mRTL3UTR-F (5' AACCATTTGCCAGCGTGTGAG 3') to amplify a ~4.4 kb band using clone *Rtl-1A* (a mouse cDNA identified in the phage library screen described previously) as a template. Since both PCR products contain the desired insert and vector sequences, construction of the desired clones involved a simple self-ligation reaction. Between 50 - 100 ng of QiaQuick PCR column purified DNA was phosphorylated in a volume of 50 µl with 6 U of T4 polynucleotide kinase (New England BioLabs Inc.) in 1x T4 DNA ligase buffer (Gibco Life Technologies) (50 mM Tris-HCl pH 7.6, 10 mM MgCl<sub>2</sub>, 1 mM ATP, 1 mM DTT, and 5% w/v polyethylene glycol-8000) for 30 min at 37°C. The reaction was purified with a QiaQuick PCR spin column and eluted in 50 µl of TE. Ligation reactions were performed with 50 ng of DNA in a 50 µl volume of 1x DNA ligase buffer with 1 U of T4 DNA ligase (Gibco Life Technologies) and incubated at 16°C overnight. Transformations were performed using 4 µl of each ligation reaction incubated with 50 µl of competent DH5α cells (Gibco Life Technologies) on ice for at least 30 min, heat

shocked at 42°C for 90 sec, placed on ice for 5 min, and spread onto LB plates supplemented with 100 µg/ml of ampicillin. The plates were inverted and incubated at 37°C overnight.

### **Plasmid DNA isolation**

Plasmids were obtained from isolated *E. coli* DH5α colonies. Plasmid DNA was purified from overnight cultures grown in LB supplemented with 100 µg/ml ampicillin using Qiaprep Spin Miniprep columns (Qiagen) and eluted in TE. Isolated plasmid DNA was analysed by digestion with appropriate restriction endonucleases and DNA sequencing.

### **DNA sequencing**

DNA sequencing was performed as suggested by the manufacturer (Pharmacia Inc.). Approximately 2 µg of plasmid DNA in a volume of 24 µl was denatured with 16 µl of 1M NaOH at room temperature for 10 min, neutralized with 7 µl of 3 M sodium acetate pH 5.3, and precipitated with 120 µl of ice-cold 100% ethanol. The DNA was pelleted at 4°C by centrifugation at 12,000 rpm for 10 min, rinsed with 200 µl of 70% ice-cold ethanol, and air-dried at room temperature. The pellet was resuspended in 11 µl of dH<sub>2</sub>O and 1 µl of primer (25 ng/µl). 2 µl of annealing buffer (1 M Tris-HCl pH 7.6, 100 mM MgCl<sub>2</sub>, and 160 mM DTT) was added followed by a 5 min incubation at 65°C, 10 min at 37°C, and 5 min at room temperature. 3 µl of labeling mix (1.375 mM of each dATP, dGTP, and dTTP, and 33.5 mM NaCl), 1 µl of α<sup>35</sup>S dCTP (Amersham Life

Science), and 2  $\mu$ l of diluted T7 DNA polymerase were added and incubated at room temperature for 5 min. The reaction was then incubated in the presence of each of the ddNTPs for 5 min and terminated with stop solution (0.3% each of bromophenol blue and Xylene cyanol FF, 10 mM EDTA pH 7.5, and 97.5% deionized formamide). All samples were denatured at 80°C prior to electrophoresis on a denaturing 6% polyacrylamide sequencing gel.

When longer DNA sequence reads were required, sequencing was performed by the DNA sequencing core facility (Canadian Genetic Diseases Network) with a LiCOR automated DNA sequencing machine.

### **Radiation hybrid mapping**

Radiation hybrid (RH) mapping was performed in collaboration with Dr. S.W. Scherer (The Hospital for Sick Children, Toronto, Canada). Briefly, primers RTL-F and RTL-R were used to amplify DNA samples from human GB3 RH panels (Whitehead Institute). The pattern of positive clones was used to determine the chromosome location.

### ***In situ* hybridization**

#### *Tissue sectioning*

Human eyes were obtained from the Canadian Eye Bank approximately 16 hours postmortem. The retina, while attached to the sclera, was removed from other structural tissues in ice cold 1x PBS (2.7 mM KCl, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>, 137 mM NaCl, and 0.5 mM Na<sub>2</sub>HPO<sub>4</sub>). Mature murine brains were dissected from adult (>6 mos) mice. For developmental expression studies, C57BL/6 mice were paired together to obtain embryos.

Female mice with plugs detected the morning following pairing were considered to be at E0.5. Embryos from E12.5, E13.5, E14.5, and E15.5 were dissected in ice cold 1x PBS. All tissues were fixed overnight in 4% paraformaldehyde (PFA), 1x PBS. Following fixation, the tissues were rinsed three times in 1x PBS at 4°C in overnight successions, followed by an overnight incubation in 30% w/v sucrose, 1x PBS at 4°C or until the tissues had settled to the bottom of the tube.

After sedimentation, the tissues were removed from the sucrose solution and submerged in O.C.T. (Tissue-Tek, Sakura Finetek USA, Inc.) and incubated for at least 30 min at room temperature. The tissues were transferred to fresh O.C.T. plastic embedding blocks. These blocks were frozen on dry ice and stored at -20°C.

The blocks were mounted with O.C.T. and 14 µm sections obtained using a Leica cryostat were placed onto Silane-Prep™ slides (Sigma Diagnostics), air dried for at least 2 hours at room temperature, and frozen at -20°C in a sealed container.

### *In situ riboprobes*

To construct antisense (AS) and sense (S) probes, 10 µg of RTLCDs was linearized with 30 U of either *Sfi*I (AS) or *Hind*III (S) in a 50 µl volume for at least 3 hours at 50°C or 37°C, respectively, in the appropriate reaction buffer. The DNA samples were digested with 20 ng of proteinase K with the addition of 50 µl of 2x proteinase K buffer (10 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, and 0.5% SDS) at 37°C for 30 min. To precipitate the DNA, 10 µl of 3M sodium acetate pH 5.3 was added and a phenol/chloroform extraction and ethanol precipitation was performed. The DNA was resuspended in 11 µl of DEPC ddH<sub>2</sub>O for *in vitro* transcription reactions using either

T3 (AS) or T7 (S) RNA polymerases. For each reaction, 1  $\mu$ g of linearized DNA template was raised to a volume of 11  $\mu$ l and an 7  $\mu$ l cocktail mix containing 1x transcription buffer (40 mM Tris-HCl, pH 8.0, 6 mM MgCl<sub>2</sub>, 10 mM DTT, 2 mM spermidine), 14.2 mM DTT, and 2.85x digoxigenin labeling mix) (Roche Diagnostics Corporation), along with 1  $\mu$ l of RNase inhibitor (Roche Diagnostics Corporation), and 20 U of either T3 or T7 (20 U/ $\mu$ l) RNA polymerase were added and incubated at 37°C for 1 hour. An additional 20 U of T3 or T7 RNA polymerase was added to their respective tubes and incubated for an additional hour at 37°C. Following *in vitro* transcription, the RNA was precipitated with 2.5  $\mu$ l of 4 M LiCl, 2  $\mu$ l of 0.2 M EDTA, and 75  $\mu$ l of ice-cold 100% ethanol. The RNA was pelleted at 4°C by centrifugation at 12,000 rpm for 10 min and resuspended in 110  $\mu$ l of DEPC ddH<sub>2</sub>O.

#### *In situ hybridization*

Prior to *in situ* hybridization, glass slide dishes (Wheaton) were baked at 220°C for 6 hours to destroy contaminating RNases. All solutions were prepared from DEPC ddH<sub>2</sub>O, and all washes were conducted in 200 ml volumes at room temperature, unless otherwise noted.

Slide containers were removed from -20°C and placed at room temperature for at least 2 hours before they were opened to avoid condensation. Cryostat tissue slides were heated in a hybridization oven at 50°C for 20 min, fixed in 4% PFA, 1x PBS for 20 min, and rinsed twice in 1x PBS for 5 min. The sections were then treated with a solution containing 20  $\mu$ g/ml proteinase K, 50 mM Tris-HCl pH 8.0 and 5 mM EDTA pH 8.0 for 10 minutes and then fixed in 4% PFA, 1x PBS for 20 minutes. The slides were washed in

DEPC ddH<sub>2</sub>O for 1 min, and then acetylated in a 200 ml solution containing 20.6  $\mu$ M acetic anhydride, 22.4 mM sodium hydroxide, and 93.6 mM triethanolamine (Sigma Diagnostics). Following fixation and acetylation, the slides were incubated in slide mailers containing ~ 8 ml of prehybridization solution (50% formamide, 5x SSC, 5x Denhardt's solution, 0.25 mg/ml yeast tRNA, and 0.5 mg/ml salmon sperm DNA) for at least 3 hours at 60°C. Approximately 25  $\mu$ l of digoxigenin-labeled sense or antisense probe (~5  $\mu$ g of riboprobe) were added to 4 ml of hybridization solution (identical to the prehybridization solution), heated to 80°C for 5 min, placed onto ice for a minimum of 3 min and added to the slide mailers. Hybridization occurred at 55°C overnight in a hybridization oven without agitation.

After hybridization, the slides were removed from the slide mailers and placed in glass slide dishes for washing (all washes were carried out at 37°C in 200 ml volumes unless stated otherwise). Slides were washed once each in 5x SSC at 55°C for 15 min, 0.2 x SSC at 55°C for 60 min, 1x RNA (0.4 M NaCl, 0.1 M Tris-HCl pH 7.5, and 0.05M EDTA) for 10 min, 1 x RNA with 20  $\mu$ g/ml RNaseA for 30 min, 1 x RNA for 5 min, 2 x SSC for 10 min, and 0.2 x SSC for 10 min.

#### *Secondary detection conditions*

The following steps were all carried out at room temperature. Washes were carried out in 200 ml volumes using glass slide dishes and 8 ml incubations were conducted in slide mailers. The slides were incubated in 200 ml of NT (0.1 M Tris-HCl pH 7.5, and 0.15 M NaCl) for 5 min. Afterwards, the slides were placed in 8 ml of 1% w/v blocking reagent (Roche Diagnostics Corporation) dissolved in NT at 37°C for 60

min. The solution was replaced with 8 ml of 1% w/v blocking solution containing a 1:2000 dilution of anti-digoxigenin alkaline phosphatase antibodies (Roche Diagnostics Corporation) and incubated for 60 min. This was followed by three 200 ml NT washes for 22 min. The slides were then incubated in 8 ml of NTMT (0.1 M NaCl, 0.1 M Tris-HCl pH 9.5, 0.05 M MgCl<sub>2</sub>, and 0.1% Tween) for 10 min, and treated with 8 ml of NTMT with 0.1% w/v levamisole (Sigma Diagnostics) for 5 min, and finally incubated in slide mailers containing 5 ml NTMT, 0.7% v/v of NBT (Roche Diagnostics Corporation) and 0.9% v/v of BCIP (Roche Diagnostics Corporation) overnight, in the dark.

### *Mounting*

Following overnight colour development, the slides were washed three times in 1x PBS at room temperature, each for 15 min, and fixed for 2 hours in 3.7% formaldehyde, MEMFA (1 M MOPS pH 7.5, 0.02 EGTA, 0.01 M MgSO<sub>4</sub>). After fixation, the slides were washed in serial for 1 min each in 200 ml of 30%, 70%, 90%, and 100% ethanol, 1:1 Xylene/ethanol, and 100% xylene. The slides were mounted with glass coverslips using 50% v/v permount (dissolved in xylene) and dried for 2-3 days in a fumehood.



## RESULTS

### **The majority of retina-only UniGene clusters represent novel genes**

To identify novel retinal-specific genes with a potential role in eye development, an *in silico* screen was performed using UniGene clusters from libraries 177 and 178 representing different size fractions of a normalized human retinal cDNA library. The two libraries collectively contain 14,422 EST entries and have been organized into 3,938 gene clusters, 348 of these represent retina-only clusters. Of these 348 entries, 203 contain two or more ESTs, and these were analyzed by BLAST. From the screen, 25 are known genes (a “known” entry defined as having the complete coding sequence available in the databases; this total excluded two different rhodopsin mRNA isoforms present in the UniGene database), 6 share moderate conservation with known proteins, and 170 have no homology with any proteins in the databases. Among the 25 known genes, 8 have been shown previously to be retina-specific or retina-abundant with an important role in eye development or maintenance of the retina.

### **Identification of a novel CUB domain-encoding gene, *RTL-1***

One UniGene cluster (Hs.60563) with three retinal ESTs (accession numbers R85884, AA013001, and AA019587), has a partial open reading frame (ORF) with ~40% amino acid identity to the CUB domains found in tolloids and neuropilins (Figure 2-1). Upon additional inspection, the partial cDNA encoded a CUB domain, and was designated as *RTL-1*, for *retinal tolloid-like*. The cDNAs for R85884 and AA013001 were ordered (Genome Systems, Inc.), however, AA013001 arrived heavily contaminated

**A**

										identity																																														
Rtl-1_C1	1	CGTWTKHA	EGGYFTSPNYP	SKYPDRE	ECVYII	IEAA-PRO	CI	E	LYF	FDEKYS	TEPSWE	CKFD	HDHIEVRDG	66	100%																																									
Np-1_C1	1	CGGTIKIEN	PGYLTS	PGYPHSYHP	SEKCEW	LIQ	AP	EPYQRI	II	IN	FNPHFD	LED	DRD-CKYD	YVEVIDG	66	40																																								
Np-2_C1	1	CGGRPN	SKDAGYITS	PGYPQ	DYPSHQ	NCEW	I	VY	AP	EPN	QKIVL	IN	FNPHFEIEKHD-CKYD	FIEIRDG	66	41																																								
Bmp-1_C2	1	CGGDVK	-KDN	GHIQS	PNYPD	DYRPSK	VC	I	WR	I	QV	SE	GFH-VGLTFQ-SFE	I	ERHDS	CAYD	YLEVRDG	64	38																																					
mTll-1_C2	1	CGGEIR	-KNE	GQIQS	PNYPD	DYRPM	KE	C	Y	W	K	I	M	V	SE	GFH-VGLTFQ-AFE	I	ERHDS	CAYD	H	LEVRDG	64	41																																	
Rtl-1_C1	67	PFGF	SPIIGR	FCGQ	ONPP	VIKSS	GR	FL	WIK	FF	AD	GE	LES	MG	F	S	A	R	Y	N	F	T	P	D	P	D	F	K	D	L	G	V	L	K	P	L	P	A	131																	
Np-1_C1	67	ENEG	R	L	W	G	K	F	C	G	K	I	A	P	S	P	V	S	S	G	P	F	L	F	I	K	F	V	S	D	Y	E	T	H	G	A	G	F	S	I	R	Y	E	I	F	K	R	G	P	E	-----	120				
Np-2_C1	67	DSE	A	D	L	L	G	K	H	C	G	N	I	A	P	P	T	I	I	S	S	G	S	V	L	I	K	F	T	S	D	Y	A	R	Q	G	A	G	F	S	L	R	Y	E	I	F	K	T	G	S	E	D	-----	121		
Bmp-1_C2	65	HSE	S	N	L	I	G	R	Y	C	G	Y	E	N	F	D	D	I	K	S	T	S	S	R	L	W	L	K	F	V	S	D	G	S	I	N	K	A	G	F	A	V	N	--	F	F	K	E	V	D	E	-----	116			
mTll-1_C2	65	A	S	E	N	S	P	L	I	G	R	F	C	G	Y	D	K	P	E	D	I	R	S	T	S	N	T	L	W	M	K	F	V	S	D	G	T	V	N	K	A	G	F	A	N	--	F	F	K	E	E	D	E	-----	116	

**B**

Rtl-1_C2	1	CE	F	E	M	G	-G	P	E	G	I	V	E	S	I	Q	I	L	K	E	G	K	A	S	A	S	E	A	V	D	C	K	W	I	R	A	P	P	R	S	K	-I	Y	L	R	F	L	-D	Y	E	M	Q	N	S	N	E	C	K	R	N	61	100%		
Np-1_C1	1	CGGT	I	K	I	E	N	P	G	Y	L	T	S	-P	G	Y	P	H	S	Y	H	P	S	E	K	---	C	E	W	L	I	Q	A	P	E	P	Y	Q	R	I	I	I	N	F	N	P	H	F	D	L	E	D	R	D	-C	K	Y	D	59	26				
Np-2_C1	1	CGGR	P	N	S	K	D	A	G	Y	I	T	S	-P	G	Y	P	Q	D	Y	P	S	H	Q	---	C	E	W	I	V	Y	A	P	E	P	N	Q	K	I	V	L	N	F	N	P	H	F	E	I	E	K	H	D	-C	K	Y	D	59	26					
Bmp-1_C2	1	CGGD	Y	K	-K	D	N	G	H	I	Q	S	-P	N	Y	P	D	D	Y	R	P	S	K	---	C	I	W	R	I	Q	V	S	E	G	F	H	-V	G	L	T	F	Q	-S	F	E	I	E	R	H	D	S	C	A	Y	D	57	23							
mTll-1_C2	1	CGGE	I	R	-K	N	E	G	Q	I	Q	S	-P	N	Y	P	D	D	Y	R	P	M	K	---	C	Y	W	K	I	M	V	S	E	G	Y	H	-V	G	L	T	F	Q	-A	F	E	I	E	R	H	D	S	C	A	Y	D	57	27							
Rtl-1_C2	62	F	V	A	Y	D	G	S	S	S	V	E	D	L	K	A	K	F	C	S	T	V	T	N	D	V	M	L	R	T	G	-L	G	V	I	R	M	W	A	D	E	G	S	R	N	S	R	F	Q	M	L	F	T	S	F	Q	E	P	P	--	120			
Np-1_C1	60	Y	V	E	I	D	G	E	N	E	G	R	L	W	G	K	F	C	G	K	I	A	P	S	P	V	S	S	G	P	F	L	F	I	K	F	V	S	D	Y	E	T	H	G	A	G	F	S	I	R	Y	E	I	F	K	R	G	P	E	-----	120			
Np-2_C1	60	F	I	E	I	R	D	G	S	E	S	A	D	L	L	G	K	H	C	G	N	I	A	P	P	T	I	I	S	S	G	S	V	L	I	K	F	T	S	D	Y	A	R	Q	G	A	G	F	S	L	R	Y	E	I	F	K	T	G	S	E	D	-----	121	
Bmp-1_C2	58	Y	L	E	V	R	D	G	H	S	E	S	N	L	I	G	R	Y	C	G	Y	E	N	F	D	D	I	K	S	T	S	R	L	W	L	K	F	V	S	D	G	S	I	N	K	A	G	F	A	V	N	--	F	F	K	E	V	D	E	-----	116			
mTll-1_C2	58	H	L	E	V	R	D	G	A	S	E	N	S	P	L	I	G	R	F	C	G	Y	D	K	P	E	D	I	R	S	T	S	N	T	L	W	M	K	F	V	S	D	G	T	V	N	K	A	G	F	A	N	--	F	F	K	E	E	D	E	-----	116		

**Figure 2-1.** Amino acid alignments between Rtl-1 CUB1 (A) and CUB2 (B) with CUB domains of murine proteins most related to Rtl-1.

Identities are highlighted in yellow and similarities are outlined in boxes. The aligned protein domains are: Rtl-1\_C1 (Rtl-1, CUB1), Rtl-1\_C2 (Rtl-1, CUB2), np-1\_C1 (neuropilin-1, CUB1), np-2\_C1 (neuropilin-2, CUB1), Bmp-1\_C2 (Bone morphogenetic protein-1, CUB2), and mTll-1\_C2 (mammalian tolloid-like-1, CUB2).

**Table 2-1**

## UniGenes representing known genes

UniGene	Gene	Reference
<b>Hs.308</b>	arrestin-C	Craft <i>et al.</i> , 1994
<b>Hs.139263</b>	L-type calcium channel alpha-1 subunit	Strom <i>et al.</i> , 1998
<b>Hs.159437</b>	Homo sapiens prospero-related homeobox 1 ( <i>PROX1</i> )	Zinovieva <i>et al.</i> , 1996
<b>Hs.232072</b>	Usher syndrome 2A	Eudy <i>et al.</i> , 1998
<b>Hs.249186</b>	cone rod homeobox ( <i>CRX</i> )	Freund <i>et al.</i> , 1997
<b>Hs.251687</b>	Retinitis pigmentosa 1 (autosomal dominant) ( <i>RPI</i> )	Pierce <i>et al.</i> , 1999
<b>Hs.267674</b>	guanylate cyclase activating protein 1 ( <i>GCAP</i> )	Subbaraya <i>et al.</i> , 1994
<b>Hs.261322</b>	rhodopsin ( <i>RHO</i> )	Nathans and Hogness, 1984
Hs.33785	EUROIMAGE 191017	Auffray <i>et al.</i> , 1999
Hs.40337	FLJ11219 fis,	Isogai <i>et al.</i> , 2000
Hs.40615	KIAA0310	Nagase <i>et al.</i> , 1997
Hs.40921	KIAA0650	Ishikawa <i>et al.</i> , 1998
Hs.62813	human ortholog of ( <i>Mlt-1</i> ); zinc finger	Tateno <i>et al.</i> , 2000
Hs.169448	<i>GTBP</i> -Alternative splice; homolog of mutS	Shiwaku <i>et al.</i> , 1997
Hs.173105	KIAA1244	Nagase <i>et al.</i> , 1999
Hs.260560	(AL096857) hypothetical protein; homology to <i>BAT2</i>	Rhodes and Huckle, 1999
Hs.260716	KIAA1116	Kikuno <i>et al.</i> , 1999
Hs.267665	Estrogen-related receptor, beta ( <i>ERRB2</i> )	Giguere <i>et al.</i> , 1988
Hs.269206	cerebroside sulfate activator protein	Dewji <i>et al.</i> , 1987
Hs.269008	choline dehydrogenase	Bugert <i>et al.</i> , 2000
Hs.269214	G-protein coupled receptor 75 ( <i>GPR75</i> )	Tartelin <i>et al.</i> , 1999
Hs.269240	neuronal cell death-related protein	Peng <i>et al.</i> , 2000
Hs.269250	diacylglycerol kinase	Ding <i>et al.</i> , 1998
Hs.269245	chondromodulin I precursor ( <i>CHM-I</i> )	Shukunami <i>et al.</i> , 1998
Hs.269247	NG,NG-dimethylarginine dimethylaminohydrolase	van Asseldonk <i>et al.</i> , 2000

**bold:** important in eye development/function

**Table 2-2**

UniGenes representing novel genes with homology to known protein(s)

UniGene	Homology to	Reference
<b>Hs.60563</b>	~40% id (over 98 a.a.) to tolloid/BMP1, CUB domain	Shimell <i>et al.</i> , 1991 Wozney <i>et al.</i> , 1989
Hs.182117	85% id (24/28 a.a.) to iduronate-2-sulfatase	Wilson <i>et al.</i> , 1993
Hs.220687	80% id (28/35 a.a.) in Gamm1 and <i>S. pombe</i> , <i>C. elegans</i> ORF	Inazu, 2000 Aert <i>et al.</i> , 2000 Miller, 1999
Hs.261251	88% id (22/25 a.a.) to unnamed HERV-H protein	Lindskog, 1997
Hs.269249	42% id (47/113 a.a.) to pregnancy zone protien precursor	Devriendt <i>et al.</i> , 1991
Hs.271684	59% id (59/99 a.a.) Bem-46-like, <i>Drosophila</i>	Ghabrial <i>et al.</i> , 1998

**bold:** contains domain/motif found in developmental proteins

with coliphage T1, and no plasmid DNA could be purified from the overnight culture. The cDNA for R85884 was sequenced but it did not resemble the EST database entry nor did it have any homology to Hs.60563. Further sequence analysis confirmed that the clone received was incorrect, therefore a different cloning strategy was used.

### Cloning of *RTL-1*

To obtain a *RTL-1* cDNA, two 3'UTR primers (RTL-F 5' CTCCCTTCATTAAATCTATTATCT 3' and RTL-R 5' TCATATATGCATATTAGTGTACTG 3') were designed from the AA013001 EST entry to amplify a 220 bp product from human genomic DNA. A single band of the expected size was obtained (data not shown), labeled, and used to probe an arrayed normalized human retinal cDNA library. One positive cDNA, clone 28R163H6, was identified after screening ~20,000 cDNAs. Sequence analysis showed that the cDNA had a ~1.8 kb insert containing a partial ORF with the CUB domain.

### **Cloning full-length mouse *Rtl-1* cDNAs**

To begin to understand the biological role of *RTL-1*, the mouse ortholog (*Rtl-1*), was cloned. A 450 bp DNA fragment containing the *RTL-1* CUB domain was used to screen mouse developing eye (P0 - P3), adult mouse retina, and mouse fetal brain phage cDNA libraries from which 12 strong positive cDNAs were identified (1 from developing eye, 6 from adult retina, and 5 from fetal brain). The cDNA inserts were liberated from the phage vector by restriction endonuclease digestion, sorted according to insert size by agarose gel electrophoresis, and sequenced. The two longest cDNAs, *Rtl-1A* (2.7 kb) and *Rtl-1B* (2.8 kb), encode different ORFs (*Rtl-1A* and *Rtl-1B*, respectively). (Figures 2-2A and 2-2B). A methionine codon in a Kozak (Kozak, 1987) context (*CGCACAGAC**ATGA***), located 12 bp immediately downstream of an in-frame stop signal, was considered to be the start codon for both cDNAs. Bases in bold match the Kozak consensus. Italicized bases deviate from the consensus sequence, but are present in other genes at a frequency >20%. A third ORF was identified by database homology searches of human genomic BAC DNA sequences and was designated *RTL-1C* (Figure 2-2C).

### **Two *Rtl-1* isoforms, *Rtl-1A* and *Rtl-1B*, are putative secreted proteins**

Nucleotide sequence comparison between *Rtl-1A* and *Rtl-1B* reveal regions of identity, and also areas which are completely different, probably due to alternative splicing. Some of these alternative splices cause a frameshift thereby generating predicted ORFs differing at their carboxy termini (Figure 2-3). The *Rtl-1A* ORF is predicted to encode a 332 amino acid protein with a putative secretory pathway signal

**Figure 2-2A-C. Amino acid and nucleotide sequences of the three *Rtl-1* cDNAs.**

**A.** The in-frame stop codon 5' to the ATG is shown in italics and bold. The start codon (bold and double underlined) in the context of the Kozak sequence (double underlined) are shown. Three repeated elements are shown underlined and bolded, and the stop codon is shown in bold and italicized with an asterisk. Inverted repeats are shown with arrows, and the sequence unique to *Rtl-1A* is highlighted in yellow. The numbers on the right indicate the nucleotides (using the first A in ATG as +1) and on the left, the corresponding amino acid residues.

```

                                TGCTTTCTTTCCCTCTGTCGCTTTCCCCCACTTCT      -95
    ATTTTACCCGTGGTTTTGCATATTTAAAATCTCTGGACTCAATCCTCTTCCCTACCACCGA      -61
    AGTCTCCCCGCTTCTAAATGGAATTAGTGGAGATCGGAGCCTCTGGTGTAACGCACAGAC      -1
ATGATCTATGGACGCAGTTTGTTCACATTATAGCAAGTTTAATCATCCTCCATTCTTCT      60
1  M I Y G R S L F H I I A S L I I L H S S
    GGAGCAACCAAGAAAGGAACAGAAAAACAAATCACCCAGAAACACAGAAGTCAGTGCAG      120
21  G A T K K G T E K Q I T P E T Q K S V Q
    TGTGGAACGTGGACAAAGCATGCAGAGGGAGGTGTCTTTACGTCTCCCAATTATCCCAGC      180
41  C G T W T K H A E G G V F T S P N Y P S
    AAATATCCCCCAGACCGAGAGTGTGTCTACATCATAGAAGCTGCCCAAGGCAGTGCATT      240
61  K Y P P D R E C V Y I I E A A P R Q C I
    GAACTTTACTTTGATGAAAAATACTCAATCGAACCATCTTGGGAGTGCAAATTTGATCAT      300
81  E L Y F D E K Y S I E P S W E C K F D H
    ATTGAAGTTCGAGATGGACCCTTTGGCTTTTCTCCAATAATTGGAAGGTCTGTGGACAA      360
101 I E V R D G P F G F S P I I G R F C G Q
    CAGAATCCACCTGTAATAAAATCTAGTGGAAAGATTCTGTGGATTAAATTTTTTGCTGAT      420
121 Q N P P V I K S S G R F L W I K F F A D
    GGCGAGCTGGAGTCTATGGGATTTTCAGCTCGATATAAATTCACACCTGATCCCGACTTT      480
141 G E L E S M G F S A R Y N F T P D P D F
    AAGGACCTTGGAGTTTTGAAACCATTGCCAGCGTGTGAGTTTGAGATGGGCGGCCCGGAA      540
161 K D L G V L K P L P A C E F E M G G P E
    GGAATTGTGGAGTCCATACAGATTCTGAAGGAAGGCAAAGCTTCTGCCAGTGAAGCCGTG      600
181 G I V E S I Q I L K E G K A S A S E A V
    GATTGCAAATGGTACATCCGAGCCCCACCACGATCCAAGATTTACTTACGTTTCTTGGAC      660
201 D C K W Y I R A P P R S K I Y L R F L D
    TATGAGATGCAGAATTCAAATGAGTGCAAAAGAACTTTGTGGCGGTCTATGACGGAAGC      720
221 Y E M Q N S N E C K R N F V A V Y D G S
    AGCTCCGTGGAGGATTTGAAGGCCAAGTTCTGCAGCACAGTGGCTAATGACGTCATGCTG      780
241 S S V E D L K A K F C S T V A N D V M L
    CGCACAGGCCTTGGGGTGATCCGCATGTGGGCAGACGAGGGTAGCCGTAACAGCAGGTTT      840
261 R T G L G V I R M W A D E G S R N S R F
CAGATGCTCTTACATCCTTTCAAGAACCTCCCTGTGAAGGCAACACATTCTTCTGCCA      900
281 Q M L F T S F Q E P P C E G N T F F C H
    AGTAACATGTGTATTAACAACACGCTAGTCTGCAATGGGCTCCAGAAGTGTATATCCC      960
301 G N M C I N N T L V C N G L Q N C V Y P
    TGGGATGAAAATCACTGTAAAGATCAAACAGCCTCGZAAAAATACGTCCAAAGGAAGTC      1020
321 W D E N H C K D Q T A S *
    AGACTTTGACCAGACCGTTTTTCCAGGAGGTGTTTGGAGCCTCCCCACTATGAGTTATGTAC      1080
    GCTCAGAGGAACTGGAGCAACAGCCGACTTTGCAGACGTGGCGGAAGACTTTGAAAATTA      1140
    CCATAAACTGAGGAGATCATCTTCCAAATGCATCCATGACCATCACTGTGGATCCCAACT      1200
    GTCCAGCGCGAAAGGCAGTCGCAGTAACCTCAGCACAAGGGACGCTTCCATTTTGGCAGA      1260
    GATACCCACACAGCCAGTCAAGCCCCTCATCCCACCCGTGAACAGAAGGAACATCCTGGT      1320
    CATGAAACACAATACTACTACAAGATGCCGCCGATGCCTGTGACATCGACGAGATTGAGGA      1380
    GGTGCCACCACAAGCCACAGGCTATCCAGACACGAAAAATCTGTCCAGCGGTTCTGCCT      1440
    CATTGGGTCTCTAAGCAAACATGAATCTGAATACAACACAATAAGGGTCTAAACGAAAAT      1500
    TCAAGAGAAGAACTATTTATACAAACATGGGGACTGTGAAAAGAAAATTCCGTAGTGAAT      1560

```

TGTGAAAAGTGGACATATTTCTAAATTCATTCCACTGCTTTATCCAACTTAAGAAGTTG	1620
CAGACATGTGTATTCCCTTCGGCAAGACTCCCTGCTGCTGCACAATGATGGTCCATTTCCCT	1680
AATTTGGTGCTGATCACCAAGTGCTCCTTAGGTTTTAAATACATTCTGAGATTTACGGAA	1740
ACTTGAAGAAGAAATTAGTTCCTGATTGAGACTACCCCAGCTTTATGTTTACTGTCGTTT	1800
TCACATTATTTATATGTTGTTTACGTGTCTTTGTTACATGATTGTAATTATGTGATTTGT	1860
GGGAAAAGAAAAAAAACCACAACCTTTGGGTGAACCTTGTGTTACATGTCCATTGTTTT	1920
CATTATATAGACTGCTTGAAAATAGTGCATTCCTGCGTGATTTTGAGCAAGCTTCAGTGA	1980
AAAATGCTAAAGTGGACCATGAAAACACCATCTAGAGGCTCAATGGACTGTACACACCTG	2040
TGTTTATTATAATTAATTTAGTAAAACATAATGAGTAAAATCTAGATTGCAATAAGACA	2100
AATCTCTCTCTCCCCCTCTCTGTCGTTATCTTTGTCCTCCGCCCTTATTGTTTCTATGTA	2160
CCCCTGCTTTCCTTTTTTGTTTTACCAAGGGAAGATACCGAGTTTCACAAATGGCTTTTT	2220
GAAATCAGGTATTTTACCATTCAAACCTTACAACAGAATTTTCGTTTGATTAAAGCACTTAG	2280
CAATATGACTGAACTGACAGTTTGAGAAAATACCCTGCATGTCAAGTACTGGATATTTAG	2340
GTTGGAAAAAATGCCCTCTTTCTTAGACACATTGTAAAAGCATGCATTCTCAAATACT	2400
GCTGGTACTGGTCCATTTCTTTGTGTGCATATGCTTGCTACTTGTAAC <b>TTTTATATAAATA</b>	2460
<b>TATATAAAATGGTATAATAATTTCC</b>	2485

**Figure 2-2B. *Rtl-1B* cDNA sequence.**

The in-frame stop codon 5' to the ATG is shown in italics and bold. The start codon (bold) in the Kozak sequence (double underlined) are shown. Three repeated elements are shown underlined and bolded, and the stop codon is shown in bold and italicized with an asterisk. Sequence which differs from the *Rtl-1A* cDNA sequence is highlighted in green. Boxed and unboxed light green shading represents sequence between A1-A2 and A3-A4, respectively, as described in Figure 2-3. A potential poly-adenylation signal is underlined. The numbers on the right indicate the nucleotides (using the first A in ATG as +1) and on the left, the corresponding amino acid residues.

```

-481
-421
-361
-301
-241
-181
-121
TGCTTTCTTTCCTCTGTCGCTTTCCCCACTC
TATTTTACCCGTGGTTTGCATATTTAAAATCTCTGGACTCAATCCTCTCCCTACTACCGA
AGTCTCCCCGCTTCTAAATGGAATTAGTGGAGATCGGAGCCTCTGGTGTTAACGCACAGAC
ATGATCTATGGACGCAGTTTGTTCACATTATAGCAAGTTTAATCATCCTCCATTCTTCT
60
1  M I Y G R S L F H I I A S L I I L H S S
GGAGCAACCAAGAAAGGAACAGAAAACAAATCACCCCAGAAACACAGAAGTCAGTGCAG 120
21  G A T K K G T E K Q I T P E T Q K S V Q
TGTGGAACGTGGACAAAGCATGCAGAGGGAGGTGTCTTTACGTCTCCAATTATCCCAGC 180
41  C G T W T K H A E G G V F T S P N Y P S
AAATATCCCCCAGACCGAGAGTGTGTCTACATCATAGAAGCTGCCCAAGGCAGTGCATT 240
61  K Y P P D R E C V Y I I E A A P R Q C I
GAACTTTACTTTGATGAAAAATACTCAATCGAACCATCTTGGGAGTGCAAATTTGATCAT 300
81  E L Y F D E K Y S I E P S W E C K F D H
ATTGAAGTTCGAGATGGACCCTTTGGCTTTTCTCCAATAATTGGAAGGTTCTGTGGACAA 360
101 I E V R D G P F G F S P I I G R F C G Q
CAGAATCCACCTGTAATAAAATCTAGTGGAAAGATTTCTGTGGATTAAATTTTGTGCTGAT 420
121 Q N P P V I K S S G R F L W I K F F A D
GGCGAGCTGGAGTCTATGGGATTTTCAGCTCGATAATAATTCACACCTGATCCCGACTTT 480
141 G E L E S M G F S A R Y N F T P D P D F
AAGGACCTTGGAGTTTGAACCATTGCCAGATCCAGCCTTTCCTGCTTTCGGCCTCCTTA
540
161 ATGATTTTGCCTGTGCAGTTTIGAGATGGGCGGCCCGGAAGGAATTGTGGAGTCCATACAGAT
600
181 L F S V *
TCTGAAGGAAGGCAAAGCTTCTGCCAGTGAAGCCGTGGATTGCAAATGGTACATCCGAGC 660
CCCACCACGATCCAAGATTTACTTACGTTTCTTGGACTATGAGATGCAGAATTCAAATGA 720
GTGCAAAAGAACTTTGTGGCGGTCTATGACGGAAGCAGCTCCGTGGAGGATTTGAAGGC 840
CAAGTTCTGCAGCACAGTGGCTAATGACGTCATGCTGCGCACAGGCCTTGGGGTGATCCG 900
CATGTGGGCAGACGAGGGTAGCCGTAACAGCAGGTTTCAGATGCTCTTCACATCCTTTCA 960
AGAACCTCCCTGTGAAGGCAACACATTCTTCTGCCACAGTAACATGTGTATTAACAACAC 1020
GCTAGTCTGCAATGGGCTCCAGAACTGTGTATATCCCTGGGATGAAAATCACTGTAAAGA 1080
AAAGAGGAAACCACCGCTGGATCAGCTCACCAACAGCCAGTGGGACCGTCAATTCGGCT 1140
CAGTTCCTGCATTCATCAATCCCTTATTCCTTTCCTCCTCCTTTCAGATCAACAGCC 1200
TCGTAAAAAATACGTCCAAAGGAAGTCAGACTTTGACCAGACCGTTTTCCAGGAGGTGTT 1260
TGAGCCTCCCCTACTATGAGTTATGTACGCTCAGAGGAACTGGAGCAACAGCCGACTTTGC 1320
AGACGTGGCGGAAGACTTTGAAAATTACCATAAACTGAGGAGATCATCTTCCAAATGCAT 1380
CCATGACCATCACTGTGGATCCCAACTGTCCAGCGCGAAAGGCAGTCGCAGTAACCTCAG 1440

```



CACAAGGGACGCTTCATTTTGGCAGAGATACCCACACAGCCAGTCAAGCCCCATCCC 1500  
ACCCGTGAACAGAAGGAACATCCTGGTCATGAAACACAACACTACTACAAGATGCCGCCGA 1560  
TGCCTGTGACATCGACGAGATTGAGGAGGTGCCACCACAAGCCACAGGCTATCCAGACA 1620  
CGAAAAATCTGTCCAGCGGTTCTGCCTCATTGGTCTCTAAGCAAACATGAATCTGAATA 1680  
CAACACAAC TAGGGTCTAAACGAAAATTCAAG 1740  
[REDACTED] 1800  
[REDACTED] 1860  
[REDACTED] 1920  
[REDACTED] 1980  
[REDACTED] 2040  
[REDACTED] 2100  
[REDACTED] 2160  
[REDACTED] 2220  
[REDACTED] 2280  
[REDACTED] 2340  
[REDACTED] 2400  
[REDACTED] 2413


**Figure 2-2C.** The putative *Rtl-1C* sequence.

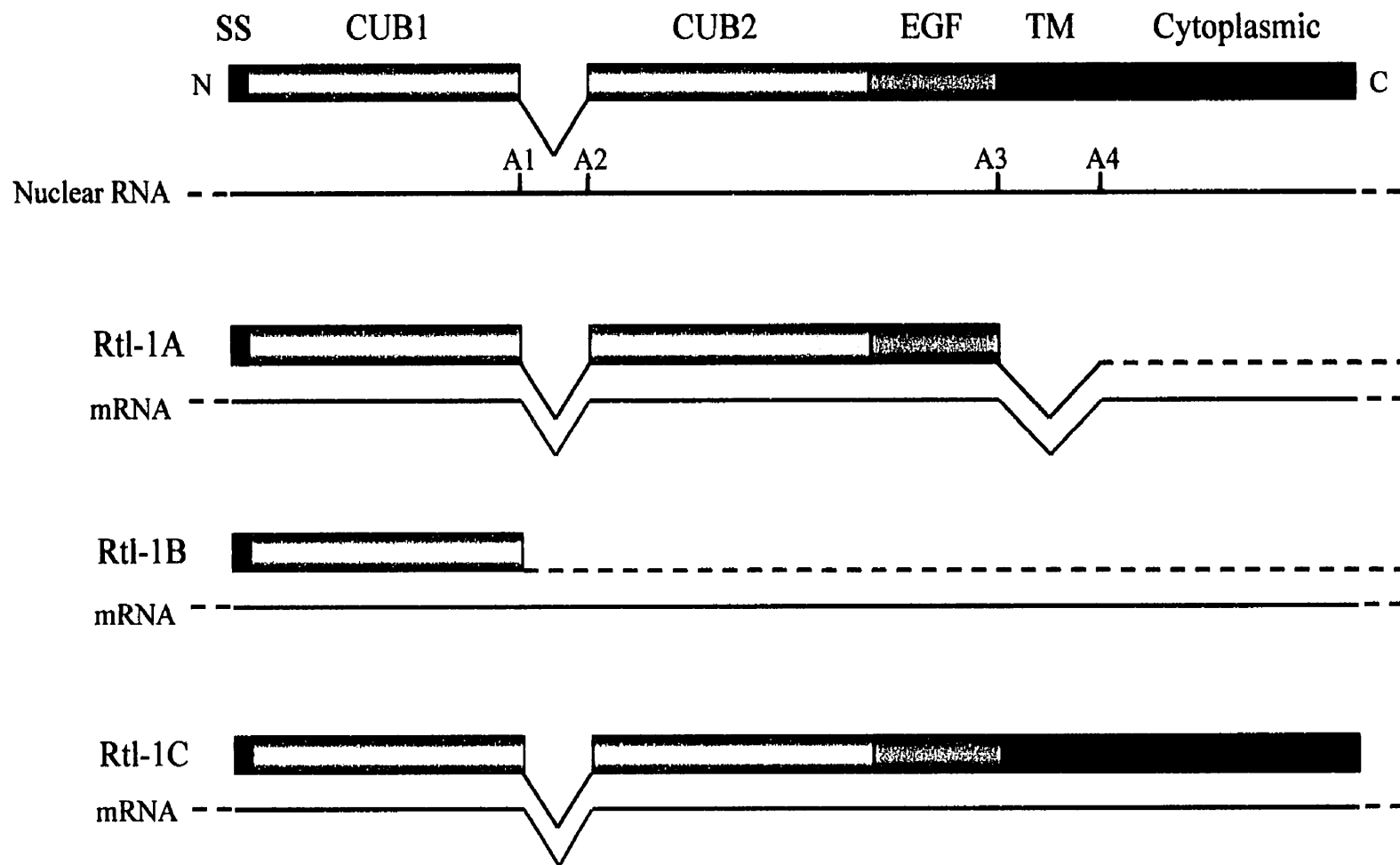
The in-frame stop codon 5' to the ATG is shown in italics and bold. The start codon (bold) in the Kozak sequence (double underlined) are shown. Three repeated elements are shown underlined and bolded, and the stop codon is shown in bold and italicized with an asterisk. Sequence which differs from the *Rtl-1A* cDNA sequence is highlighted in green. A potential polyadenylation signal is underlined.

```

-481
-421
-361
-301
-241
-181
-121
  TGCTTTCTTTCCTCTGTCGCTTCCCCACTC
TATTTTACCCGTGGTTTGCATATTTAAAATCTCTGGACTCAATCCTCTCCCTACTACCGA
AGTCTCCCCGCTTCTAAATGGAATTAGTGGAGATCGGAGCCTCTGGTGTAACGCACAGAC
ATGATCTATGGACGCAGTTTGTTCACATTATAGCAAGTTTAATCATCCTCCATTCTTCT
  60
 1  M I Y G R S L F H I I A S L I I L H S S
   GGAGCAACCAAGAAAGGAACAGAAAACAAATCACCCCAGAAACACAGAAGTCAGTGCAG
21  G A T K K G T E K Q I T P E T Q K S V Q
   TGTGGAACGTGGACAAAGCATGCAGAGGGAGGTGTCTTTACGTCTCCCAATTATCCCAGC
41  C G T W T K H A E G G V F T S P N Y P S
   AAATATCCCCCAGACCGAGAGTGTGTCTACATCATAGAAGCTGCCCAAGGCAGTGCATT
61  K Y P P D R E C V Y I I E A A P R Q C I
   GAACTTTACTTTGATGAAAAATACTCAATCGAACCATCTTGGGAGTGCAAATTTGATCAT
81  E L Y F D E K Y S I E P S W E C K F D H
   ATTGAAGTTCGAGATGGACCCTTTGGCTTTTCTCCAATAATTGGAAGGTTCTGTGGACAA
101 I E V R D G P F G F S P I I G R F C G Q
   CAGAATCCACCTGTAATAAAATCTAGTGGAAAGATTTCTGTGGATTAATTTTTTGTGAT
121 Q N P P V I K S S G R F L W I K F F A D
   GCGGAGCTGGAGTCTATGGGATTTTTCAGCTCGATATAATTTACACCTGATCCCGACTTT
141 G E L E S M G F S A R Y N F T P D P D F
   AAGGACCTTGGAGTTTGAAGCCATTGCCAGCGTGTGAGTTTGAGATGGGCGGCCCGGAA
161 K D L G V L K P L P A C E F E M G G P E
   GGAATTGTGGAGTCCATACAGATTCTGAAGGAAGGCAAAGCTTCTGCCAGTGAAGCCGTG
181 G I V E S I Q I L K E G K A S A S E A V
   GATTGCAAATGGTACATCCGAGCCCCACCACGATCCAAGATTTACTTACGTTTCTTGGAC
201 D C K W Y I R A P P R S K I Y L R F L D
   TATGAGATGCAGAATTCAAATGAGTGCAAAAGAAACTTTGTGGCGGTCTATGACGGAAGC
221 Y E M Q N S N E C K R N F V A V Y D G S
   AGCTCCGTGGAGGATTTGAAGGCCAAGTTCTGCAGCACAGTGGCTAATGACGTCATGCTG
241 S S V E D L K A K F C S T V A N D V M L
   CGCACAGGCCTTGGGGTGTATCCGCATGTGGGCAGACGAGGGTAGCCGTAACAGCAGGTTT
261 R T G L G V I R M W A D E G S R N S R F
   CAGATGCTCTTTCACATCCTTTCAAGAACCTCCCTGTGAAGGCAACACATTCTTCTGCCAC
281 Q M L F T S F Q E P P C E G N T F F C H
   AGTAACATGTGTATTAACAACACGCTAGTCTGCAATGGGCTCCAGAACTGTGTATATCCC
301 S N M C I N N T L V C N G L Q N C V Y P
   TGGGATGAAAATCACTGTAAAGAAAAGAGGAAAACCAGCCTGCTGGATCAGCTGACCAAC
321 W D E N H C K E K R K T S L L D Q L T N
   ACCAGTGGGACGGTCATTGGCGTGACTTCTGCATTGTGATCATCCTTATTATCGTTTCT
  1140

```

341	T S G T V I G V T S C I V I I L I I V S	
	GTCATTGTACAGATCAAACAGCCTCGTAAAAAATACGTCCAAAGGAAGTCAGACTTTGAC	1200
361	V I V Q I K Q P R K K Y V Q R K S D F D	
	CAGACCGTTTTCCAGGAGGTGTTTGAGCCTCCCACTATGAGTTATGTACGCTCAGAGGA	1260
381	Q T V F Q E V F E P P H Y E L C T L R G	
	ACTGGAGCAACAGCCGACTTTGCAGACGTGGCGGAAGACTTTGAAAATTACCATAAACTG	1320
401	T G A T A D F A D V A E D F E N Y H K L	
	AGGAGATCATCTTCCAAATGCATCCATGACCATCACTGTGGATCCCAACTGTCCAGCGCG	1380
421	R R S S S K C I H D H H C G S Q L S S A	
	AAAGGCAGTCGCAGTAACCTCAGCACAAGGGACGCTTCCATTTGGCAGAGATAACCCACA	1440
441	K G S R S N L S T R D A S I L A E I P T	
	CAGCCAGTCAAGCCCCTCATCCCACCCGTGAACAGAAGGAACATCCTGGTCATGAAACAC	1500
461	Q P V K P L I P P V N R R N I L V M K H	
	AACTACTCACAAGATGCCGCCGATGCCTGTGACATCGACGAGATTGAGGAGGTGCCCCACC	1560
481	N Y S Q D A A D A C D I D E I E E V P T	
	ACAAGCCACAGGCTATCCAGACACGAAAAATCTGTCCAGCGGTTCTGCCTCATTGGGTCT	1620
501	T S H R L S R H E K S V Q R F C L I G S	
	CTAAGCAAACATGAATCTGAATACAACAACACTAGGGTCTAAACGAAAATTCAAG	1680
521	L S K H E S E Y N T T R V *	
		1740
		1800
		1860
		1920
		1980
		2040
		2100
		2160
		2220
		2280
		2340
		2376



58

**Figure 2-3.** Differential usage of splice acceptor sites generate different Rtl-1 isoforms.

Shaded boxes represent ORF sequence while broken lines indicate UTR sequence, respectively. A1-4 are the relevant splice acceptor sites used by each of the Rtl-1 isoforms. (See Figure 2-2B for sequences between A1-2 and A3-4.)

sequence, two consecutive CUB domains (CUB1 and CUB2), and a single EGF-like motif. *Rtl-1B*, on the other hand, has a predicted ORF of 184 amino acids with a putative signal sequence and a single CUB domain (CUB1) (Figure 1-5).

Using *Rtl-1A* and *Rtl-1B* nucleotide sequences to query against dbEST, 58 human ESTs with ~75% nucleotide identity to the *Rtl-1* cDNAs were identified. The overlapping ESTs, clustered together as UniGene Hs. 6283, were aligned to generate a 465 amino acid partial ORF (designated RTL-2) encoding two CUB domains, an EGF-like motif, a putative transmembrane (TM) region, and a cytoplasmic tail.

Additional nucleotide sequence alignments between *Rtl-1B* and *RTL-2* revealed a 1,080 bp stretch of *Rtl-1B* 3'UTR sequence with ~75% identity to the nucleotide sequence encoding the CUB2, EGF-like motif, and putative TM and cytoplasmic regions of *RTL-2*. However, analysis of multiple sequence reactions from both strands of the *Rtl-1B* cDNA and mouse genomic DNA show a stop codon immediately after CUB1 introduced by a splice acceptor site that changes the downstream reading frame (data not shown).

The *Rtl-1A* cDNA uses a different splice acceptor site which does not alter the reading frame and thus also encodes the second CUB domain and the EGF-like motif. Although the cDNA lacks the nucleotide sequence encoding the the putative TM domain, it has sequence encoding the cytoplasmic region. However, a stop codon immediately following the EGF-like motif is expected to prevent translation into the cytoplasmic region.

### **Human RTL-1C and murine Rtl-1C are probable Rtl-1 isoforms**

From the original normalized human retinal cDNA library screen, a single *RTL-1* cDNA with a partial ORF encoding CUB1 was identified. By examining the homologies between murine *Rtl-1A* and *Rtl-1B* with human *RTL-2*, and the different splice acceptor sites, a large 533 amino acid ORF (Rtl-1C) (Figure 2-2C) with the same domain organization as *RTL-2* was constructed. To determine the full coding sequence of human *RTL-1*, the mouse *Rtl-1C* nucleotide sequence was used with BLAST to query human genomic DNA sequences in the *high throughput genomic sequences* (htgs) database to identify human *RTL-1* exons. These exons were “spliced” together to form a hypothetical cDNA sequence. The 532 amino acid ORF with the same domain organization as Rtl-1C and was therefore referred to as *RTL-1C*. *RTL-1C* shares 92% nucleotide identity and 94% amino acid identity with *Rtl-1C*. In addition, *RTL-1C* is one amino acid residue shorter than Rtl-1C - the difference occurring within the putative secretory signal sequence where the two sequences are least conserved (Figure 2-4).

### ***Rtl-1C* is the most abundant splice variant in the adult mouse retina**

To determine the relative abundance of *Rtl-1A* and *Rtl-1B/Rtl-1C* transcripts, primers mRTLCDS-F2 (5' CAGATGCTCTTCACATCC 3') and mRTL3UTR-R5 (5' GTCTGACTTGGTTTGGACG 3') which flank the putative TM-encoding region, were used to amplify 180 bp (*Rtl-1A*) and 300 bp (*Rtl-1B/Rtl-1C*) products from mouse retinal cDNA. (*Rtl-1A* transcripts lack the putative TM sequence, while *Rtl-1B/C* transcripts both retain this sequence). The 300 bp product was ~5-10x more intense than the 180 bp

```

Rtl1-1C MIYGRSLFHIIASLIILHSSGATKKGTEKQITPETQKSVQCGTWTKHAEGGVFTSPNYPSKYPPDRECVYIIEAAPRQCIELYFDEKYSIEPSWECKFDH 100
RTL-1C ML-AEPPFKVVASLIILHLSGATKKGTEKQTTSETQKSVQCGTWTKHAEGGIFTSPNYPSKYPPDRECIYIIEAAPRQCIELYFDEKYSIEPSWECKFDH 99
* . . . . .***** ***** * .***** .***** .***** .***** .***** .***** .***** .*****

Rtl1-1C IEVRDGPFGFSPIIGRFCGQONPPVIKSSGRFLWIKFFADGELESMGFSARYNFTPDPDFKDLGVLKPLPA 200
RTL-1C IEVRDGPFGFSPIIGRFCGQONPPVIKSSGRFLWIKFFADGELESMGFSARYNFTPDPDFKDLGALKPLPA 199
***** .***** .***** .***** .***** .***** .***** .***** .*****

Rtl1-1C [REDACTED] CEGNTFFCH 300
RTL-1C [REDACTED] CEGNTFFCH 299
***** .***** .***** .***** .***** .***** .***** .***** .*****

Rtl1-1C SNMCINNTLVCNGLQNCVYPWDENHCCKRRTSLLDQLTNTSGTVIGVTSCIVILIIIVSVIVQIKQPRKKYVQRKSDFDQTVFQEVFEPPHYELCTLRG 400
RTL-1C SNMCINNTLVCNGLQNCVYPWDENHCCKRRTSLLDQLTNTSGTVIGVTSCIVILIIISVIVQIKQPRKKYVQRKSDFDQTVFQEVFEPPHYELCTLRG 399
***** .***** .***** .***** .***** .***** .***** .***** .*****

Rtl1-1C TGATADFADVAEDFENYHKLRRSSSKCIHDHHCQSLSAKGSRNLSTRDASILAEIPTQPVKPLIPPVNRNRLVMKHNYSQDAADACDI DEIEEVPT 500
RTL-1C TGATADFADVADDFENYHKLRRSSSKCIHDHHCQSLSSTKGSRSNLSTRDASILTEMPTQPGKPLIPPMNRNRLVMKHSYSQDAADACDI DEIEEVPT 499
***** .***** .***** .***** .***** .***** .***** .***** .*****

Rtl1-1C TSHRLSRHEKSVQRFCLIGSLSKHESEYNTTRV 533
RTL-1C TSHRLSRHDKAVQRFCLIGSLSKHESEYNTTRV 532
***** .***** .***** .***** .***** .***** .***** .***** .*****

```

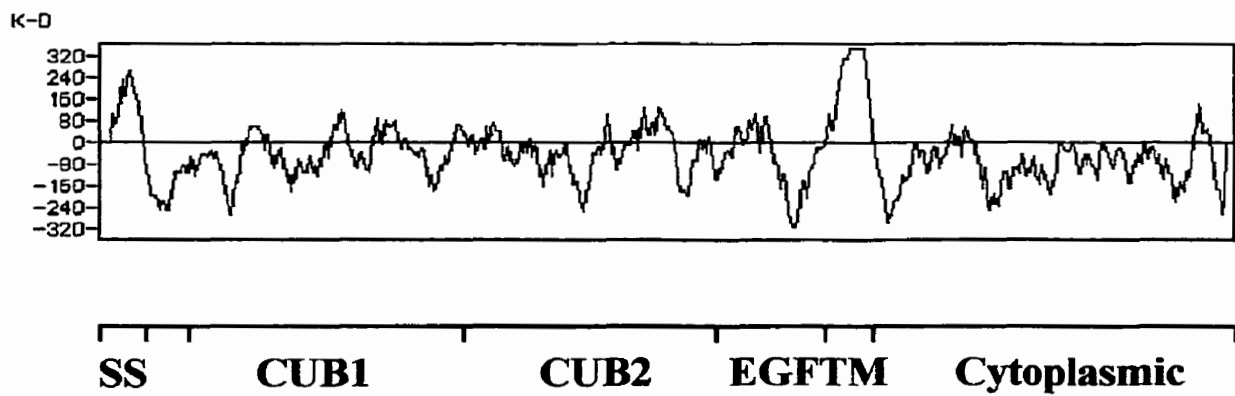
**Figure 2-4.** Amino acid alignment between Rtl-1C and RTL-1C.

Identical amino acids are indicated with an asterisk below, and similar residues are indicated with a period. Gaps are represented by a hyphen. The predicted signal sequence is boxed in white. CUB1 and CUB2 sequences are shaded yellow and green, respectively. The EGF-like motif is shaded peach, and the putative transmembrane region is coloured magenta. Conserved cysteine residues are coloured purple (in the CUB domains) or orange (outside the CUB domains), and the potential PDZ C-terminal binding site is highlighted blue.

product (Figure 2-6A), suggesting that *Rtl-1B/Rtl-1C* mRNA is present at higher levels in the adult retina than *Rtl-1A* mRNA.

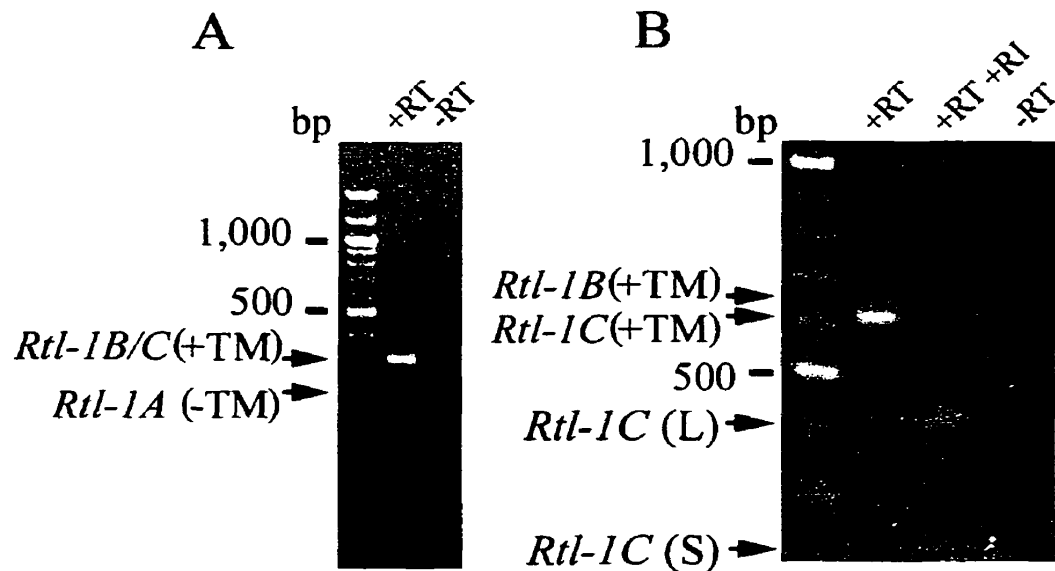
Although a cDNA encoding a *Rtl-1C* ORF has not yet been physically cloned, evidence provided by RT-PCR suggests the transcript for this predicted isoform likely exists *in vivo*. To distinguish between the two splice variants, RT-PCR was performed using primers mRTLCD5-F0 (5' ATCCCGACTTTAAGGACCTTG 3'), located immediately 5' of the splice acceptor junction site used in *Rtl-1B* and the one predicted to be used by *Rtl-1C* (the same splice acceptor site used in *Rtl-1A*), and mRTLTM-R (5' ACGATAATAAGGATGATCAC 3') which is specific for the putative TM encoding region. The expected product from a *Rtl-1B* transcript is 660 bp. For *Rtl-1C*, which is predicted to use a more 3' splice acceptor site, a 620 bp product is anticipated. Two bands of ~660 and ~620 bp were detected, the latter band being ~5x more intense (Figure 2-6B). Both products contain an internal *EcoRI* site, and this site was cleaved to verify that the products were *Rtl-1* specific bands, which yielded the expected ~410 bp and ~210 bp bands (Figure 2-6B). The larger fragments from the 660 bp product were not detected visually due to the small amount of the initial product. No product was detected with the -RT template. These data suggest the existence of *Rtl-1C* transcripts *in vivo* and at a higher level than *Rtl-1B* transcripts, at least in the adult mouse retina. No RT-PCR product from *Rtl-1A* was produced because its transcript lacks the TM-encoding region.





**Figure 2-5.** Kyte-Doolittle hydropathy plot of Rtl-1C.

The putative signal sequence is shown as a short hydrophobic region and the potential TM is represented as a large hydrophobic peak between the EGF and cytoplasmic domains. The profiles were plotted as the mean hydrophobic index versus amino acid number, using a sliding window of 19 amino acids with a 1 residue interval.



**Figure 2-6.** Detection of different *Rtl-1* mRNA splice variants in retina by RT-PCR.

(A) RT-PCR from adult mouse retinal RNA using primers mRTLCDS-F2 and mRTL3UTR-R5, which flank the putative TM domain, to distinguish and determine the relative abundance of *Rtl-1A* and *Rtl-1B/C*. No product was detected in the -RT lane.

(B) Primers mRTLCDS-F0 and mRTLTM-R were used to detect *Rtl-1B* and *Rtl-1C* splice variants. The amplified products contain an internal *EcoRI* site, which was cleaved (shown in the +RT+RI lane) to yield the desired size products. No product was detected in the -RT lane. *Rtl-1C(L)*, large fragment; *Rtl-1C(S)*, small fragment.

### **Tolloids and neuropilins share homology with Rtl-1**

All three isoforms of Rtl-1 have a predicted secretory signal sequence. The largest isoform, Rtl-1C, has two CUB domains, an EGF-like motif, a putative TM spanning region, and a cytoplasmic tail. CUB1, found in Rtl-1A, -1B, and -1C, is most related to the CUB domains in tolloids and neuropilins, sharing ~40% amino acid identity. CUB2 is more divergent than CUB1, but is also most related to the CUB domains of these two families of proteins, with ~25% amino acid identity. All four conserved cysteine residues located in CUB domains are present in both CUB1 and CUB2 (Figure 2-1). The EGF-like motif in Rtl-1A and Rtl-1C has the six conserved cysteine residues required for maintaining the topological fold. A putative transmembrane spanning region (Figure 2-5) located towards the carboxy terminus of Rtl-1C, predicted by both Kyte-Doolittle hydropathy plots (Kyte and Doolittle, 1982) and TM-Finder (Liu and Deber, 1999) is followed by a 165 a.a. cytoplasmic tail (Figure 2-4).

### **Rtl-1 has potential sites of protein modification**

To determine whether the Rtl-1 polypeptide contains potential sites of modification or regulation, the predicted protein sequences were examined using Prosite, a computer program which uses a collection of protein prediction algorithms (Hofmann *et al.*, 1999) and SignalP which identifies potential secretory signal sequences (Nielsen *et al.*, 1997). Rtl-1A, -1B, and -1C contain a potential secretory signal sequence and a possible cleavage site at residue thr22. The putative cytoplasmic region in Rtl-1C has several potential phosphorylation sites for cAMP- and cGMP-dependent protein kinases (421-424 R-R-S-S), casein kinase II (377-380 S-D-F-D, 448-451 S-T-R-D, 506-509 S-R-

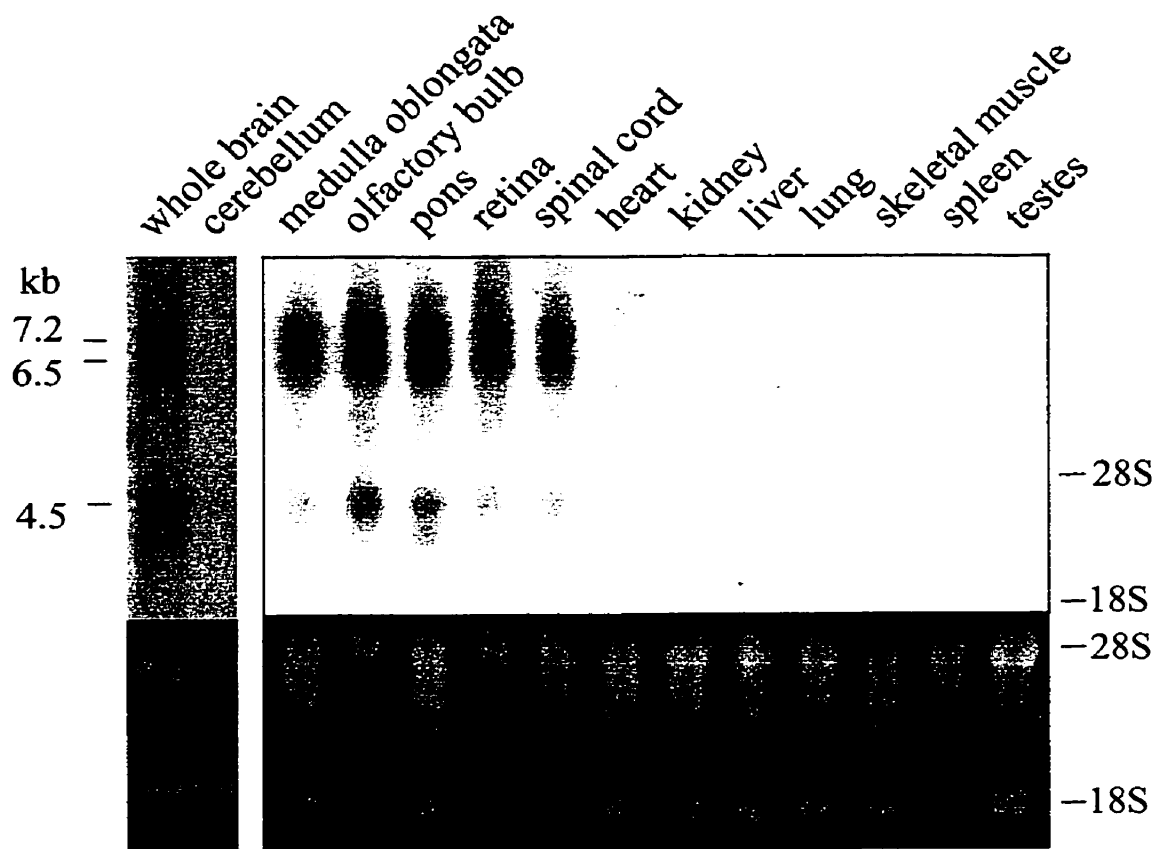
H-E, 522-525 S-K-H-E), and protein kinase C (397-399 T-L-R, 424-426 S-S-K, 439-441 S-A-K, 448-450 S-T-R, 502-504 S-H-R, 530-532 T-T-R). In addition, the last three amino acid residues at the C-terminus (T-R-V-COOH), match the consensus PDZ domain binding sequence, (S/T)-X-V-COOH (Songyang *et al.*, 1997).

### ***Rtl-1* is neurospecific**

To examine the expression pattern of *Rtl-1*, adult mouse multi-tissue northern blots were probed using a radiolabeled *Rtl-1* DNA fragment encoding CUB1. At least three bands (7.2, 6.5, and 4.5 kb), were detected in all the neural tissues examined (retina, olfactory bulb, cerebral cortex, pons, medulla, and spinal cord) with the exception of cerebellum, where no transcripts were detected (Figure 2-7). There were no bands detected in the non-neural tissues examined (lung, liver, heart, spleen, testes, kidney, and skeletal muscle) indicating that *Rtl-1* was neural specific rather than retina-specific, as originally suggested by the UniGene cluster. Identical results were obtained when DNA fragments encoding (1) CUB2 and EGF-like motifs and (2) the 3'UTR of *Rtl-1A* (which contains sequences encoding the cytoplasmic domain found in *Rtl-1C*) (data not shown).

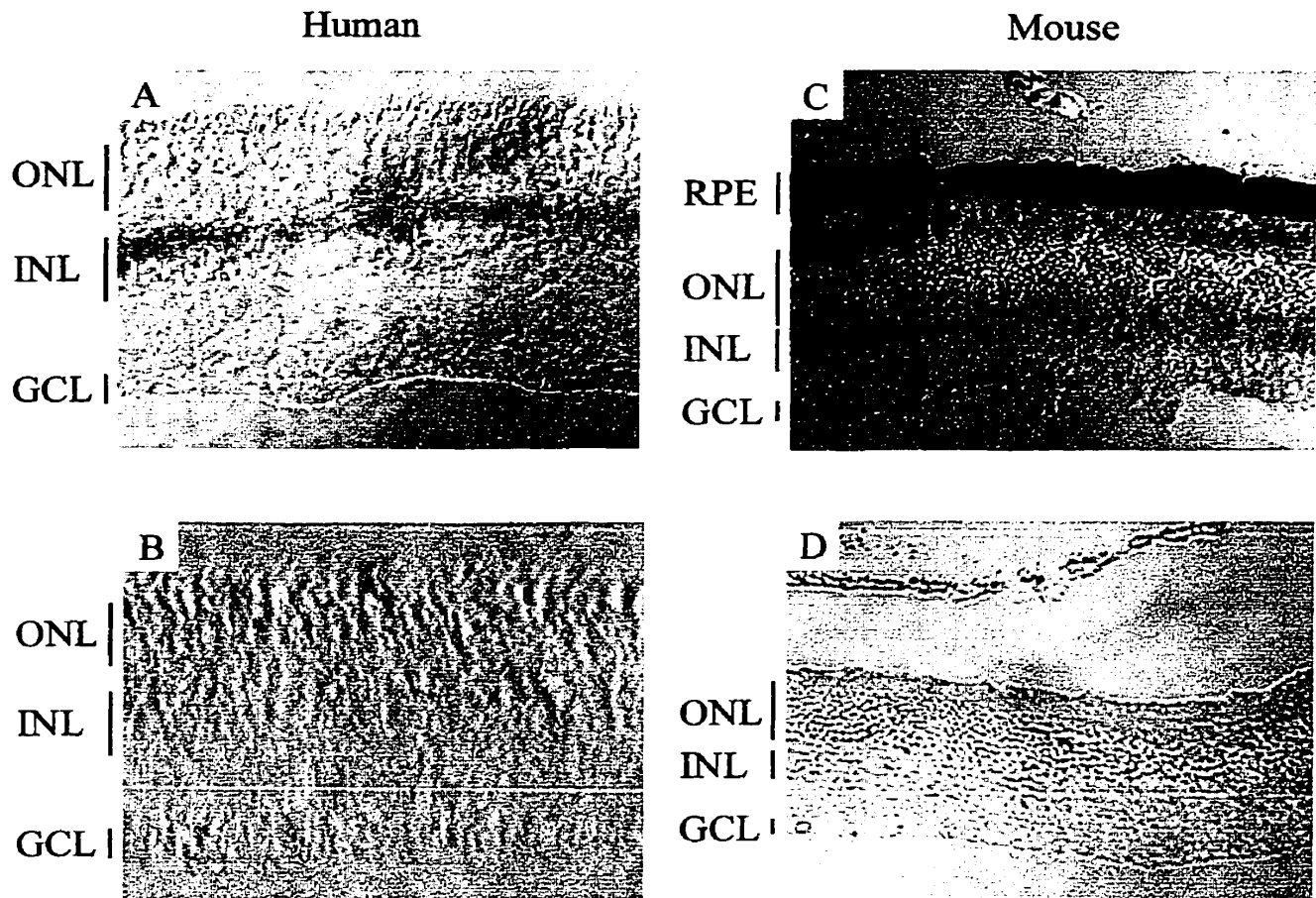
### **Expression of *Rtl-1* in mature retinal bipolar cells**

To determine the retinal expression pattern of *Rtl-1*, *in situ* hybridization was performed with adult human and mouse retinal sections using digoxigenin-labeled riboprobes corresponding to the DNA sequence encoding CUB1 (this probe was used for all subsequent *in situs*). The pattern of expression (seen as a purple blue stain under brightfield or white under darkfield illumination) in both species using the antisense (AS)



**Figure 2-7.** *Rtl-1* is neurospecific.

An adult mouse multi-tissue northern blot probed with a DNA fragment encoding CUB1 of *Rtl-1*. Approximately 10  $\mu$ g of total RNA was loaded in each lane except lanes 1 and 2, which contained  $\sim$ 5  $\mu$ g of total RNA. Three transcripts of  $\sim$ 7.2, 6.5, and 4.5 kb were detected in all neural tissues examined except in cerebellum.



**Figure 2-8.** *RTL-1* expression in the adult retina of human (A, B) and mouse (C,D).

(A, C) Antisense probe shows *RTL-1* expression (purple staining) restricted to the INL. (B, D) No staining was observed with the sense probe. ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer; RPE, retinal pigment epithelium.

probe indicated that *Rtl-1* is detectable only in the outer portion of the INL where bipolar cell nuclei reside. No specific signal was detected using the sense probe (Figure 2-8).

### ***Rtl-1* is expressed in the cerebral cortex and hippocampus**

By examining coronal sections through adult mouse brain, strong *Rtl-1* expression was observed in the CA1 and CA3 fields and the dentate gyrus of the hippocampus (Figure 2-9). In the cerebral cortex, punctate staining was observed throughout the cerebral cortical layer, however, stronger staining was observed in the region of the entorhinal cortex. *In situ* hybridization of adult mouse cerebellar sections did not detect *Rtl-1* expression in this tissue (Figure 3-2) which corroborates with the northern blot result.

### **Expression of *Rtl-1* in the developing eye**

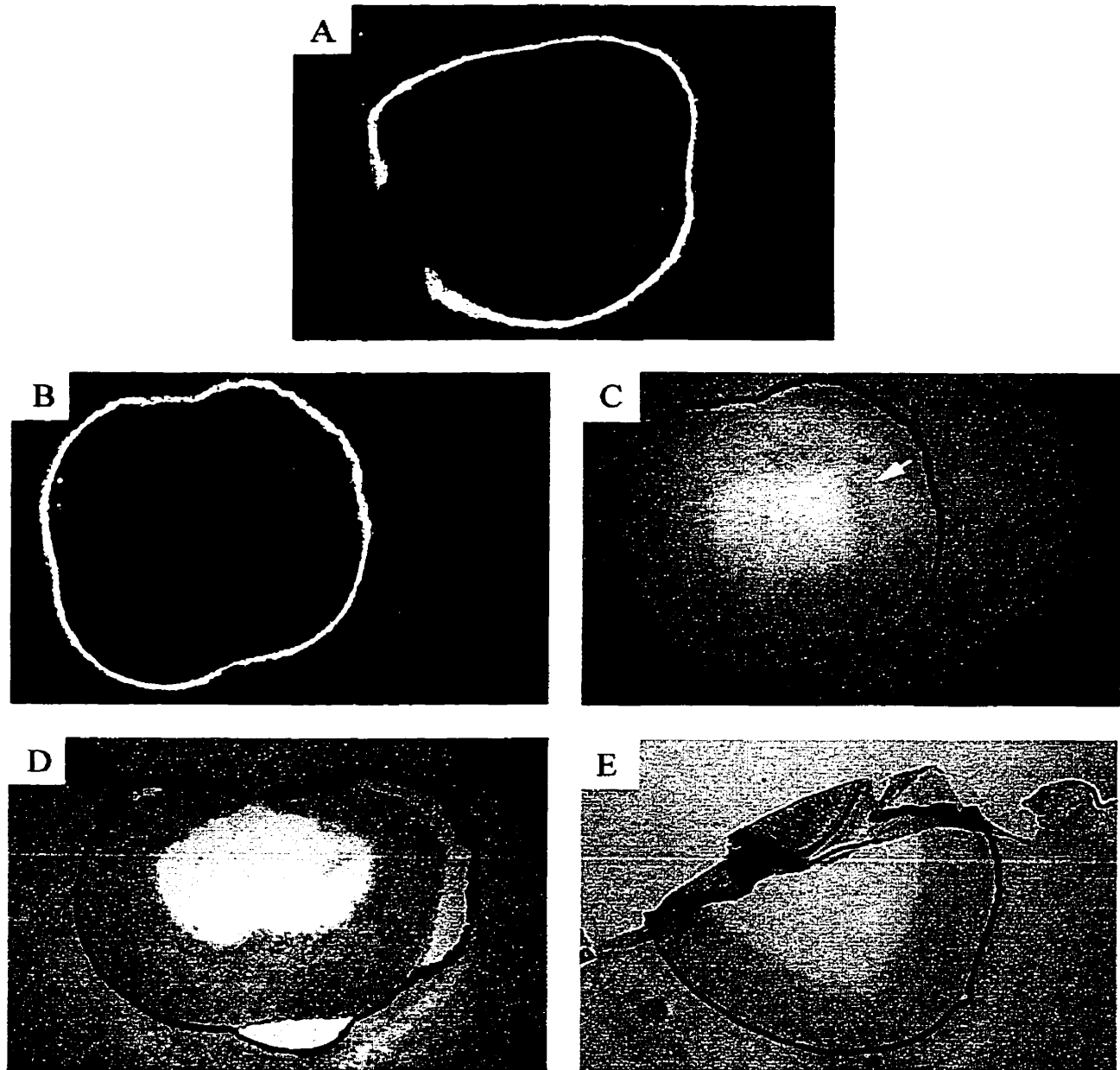
The developmental expression pattern of *Rtl-1* in mouse embryos staged between E12.5 to E15.5 was examined using sagittal and transverse sections. At E13.5, *Rtl-1* was first detected in the inner region of presumptive neuroretina where the committed and partially differentiated RGCs are found (Figure 2-10B and C). The signal is relatively weak, and is most easily detected under darkfield illumination. At E14.5 (Figure 2-10D), expression in the presumptive GCL becomes stronger. By E15.5, the expression of *Rtl-1* in the GCL decreases, but becomes detectable in the neuroblastic layer (Figure 2-10E).



**Figure 2-9.** *Rtl-1* expression in the adult mouse brain.

*Rtl-1* is expressed in the CA1, CA3, and dentate gyrus (DG) of the hippocampus, and in the cerebral cortex (C). (EC) entorhinal cortex.





**Figure 2-10.** Expression pattern of *Rtl-1* in the developing mouse eye.

*In situ* hybridization of 14  $\mu\text{m}$  eye sections using a digoxigenin-labeled riboprobe encoding CUB1 of *Rtl-1*. A) E13.5 eye. No specific signal is detected by the sense probe. B) E13.5 eye. A faint signal is detected by the antisense probe in the inner region of the presumptive neuroretina. C) E13.5 eye. The same section as in (B) but viewed under brightfield illumination. The signal is indicated by the white arrow. D) E14.5 eye. *Rtl-1* is detected in the presumptive GCL. Weak signal can be observed in the neuroblastic layer. E) E15.5 eye. *Rtl-1* expression is observed throughout the presumptive neuroretina. Under darkfield illumination (A and B), the RPE appears white and does not represent staining.

### ***Rtl-1* is expressed in the developing nervous system**

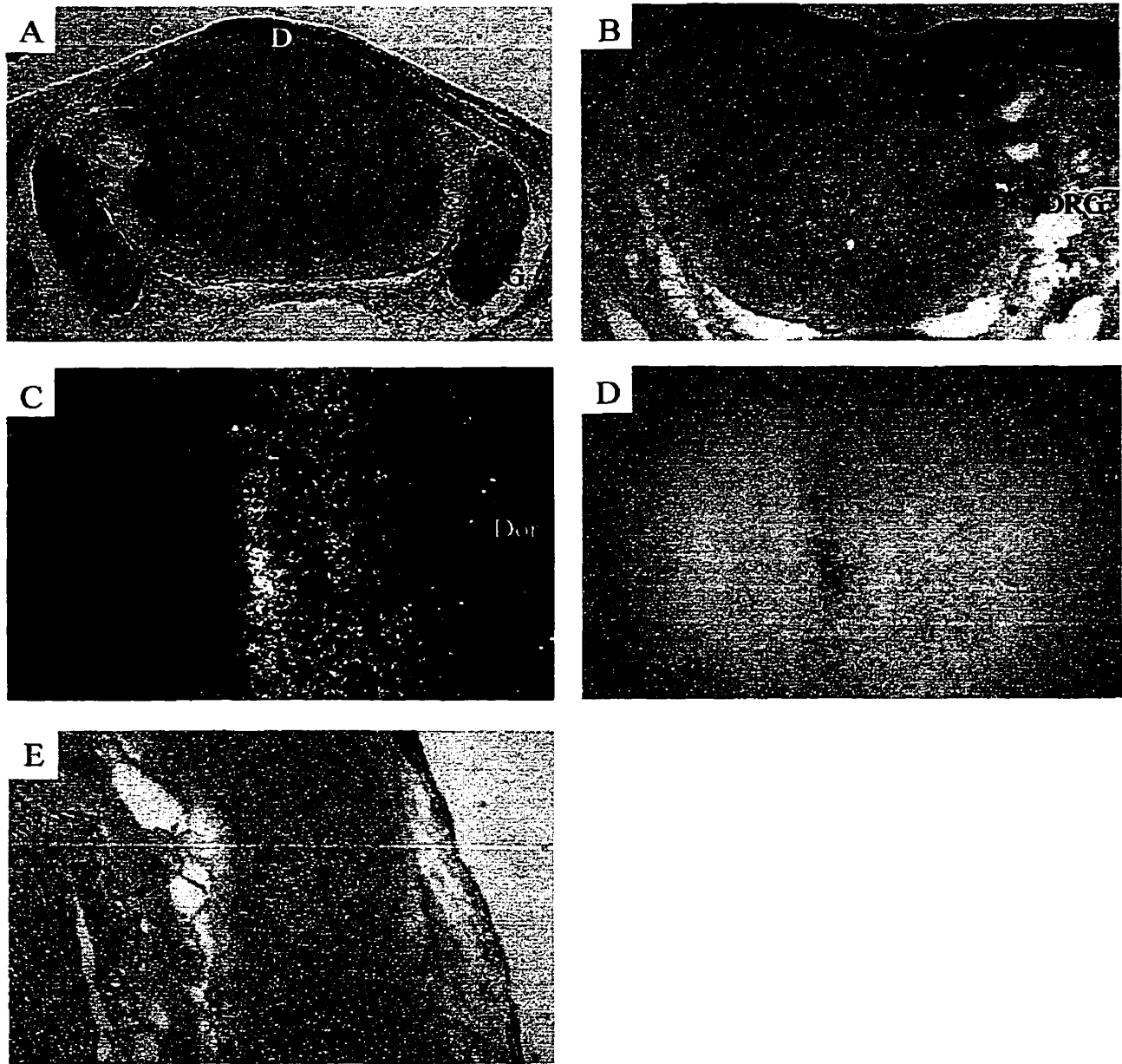
In the developing spinal cord and DRG, *Rtl-1* is expressed along the dorsoventral axis as early as E13.5 (Figure 2-11A, C and D). By examining transverse sections at the cervical level, a region of strong *Rtl-1* expression is detected in the ventral horns, near clusters of motoneuron nuclei (Figure 2-11A). Around the thoracic/lumbar level of the spinal cord, sagittal sections reveal higher expression of *Rtl-1* along the ventral spinal cord as compared to the dorsal spinal cord. (Figure 2-11C, and D). At E15.5, *Rtl-1* assumes a more uniform expression pattern (Figure 2-11B and E) in this region.

In the developing brain, *Rtl-1* was first weakly detected at E12.5 in the developing hypothalamus in two symmetrical regions adjacent to the third ventricle near the presumptive optic chiasm (Figure 2-12B). By E15.5, this expression broadens to encompass the entire region of the hypothalamus and optic chiasm (Figure 2-12C). At E13.5, expression of *Rtl-1* is also evident in the medula oblongata (Figure 2-12A).

*Rtl-1* is also expressed in other regions of neural development outside of the brain and spinal cord. For example, between E13.5 to at least E15.5, *Rtl-1* is expressed in the vomeronasal organ (Figure 2-12D) and olfactory epithelium (Figure 2-12C). At E15.5, there is also *Rtl-1* expression in the cochlear (VIII) ganglia (Figure 2-12E), trigeminal (V) ganglia (Figure 2-12C), facial (VII) ganglia (Figure 2-12C), and the glossopharyngeal (IX) ganglia (Figure 2-12E).

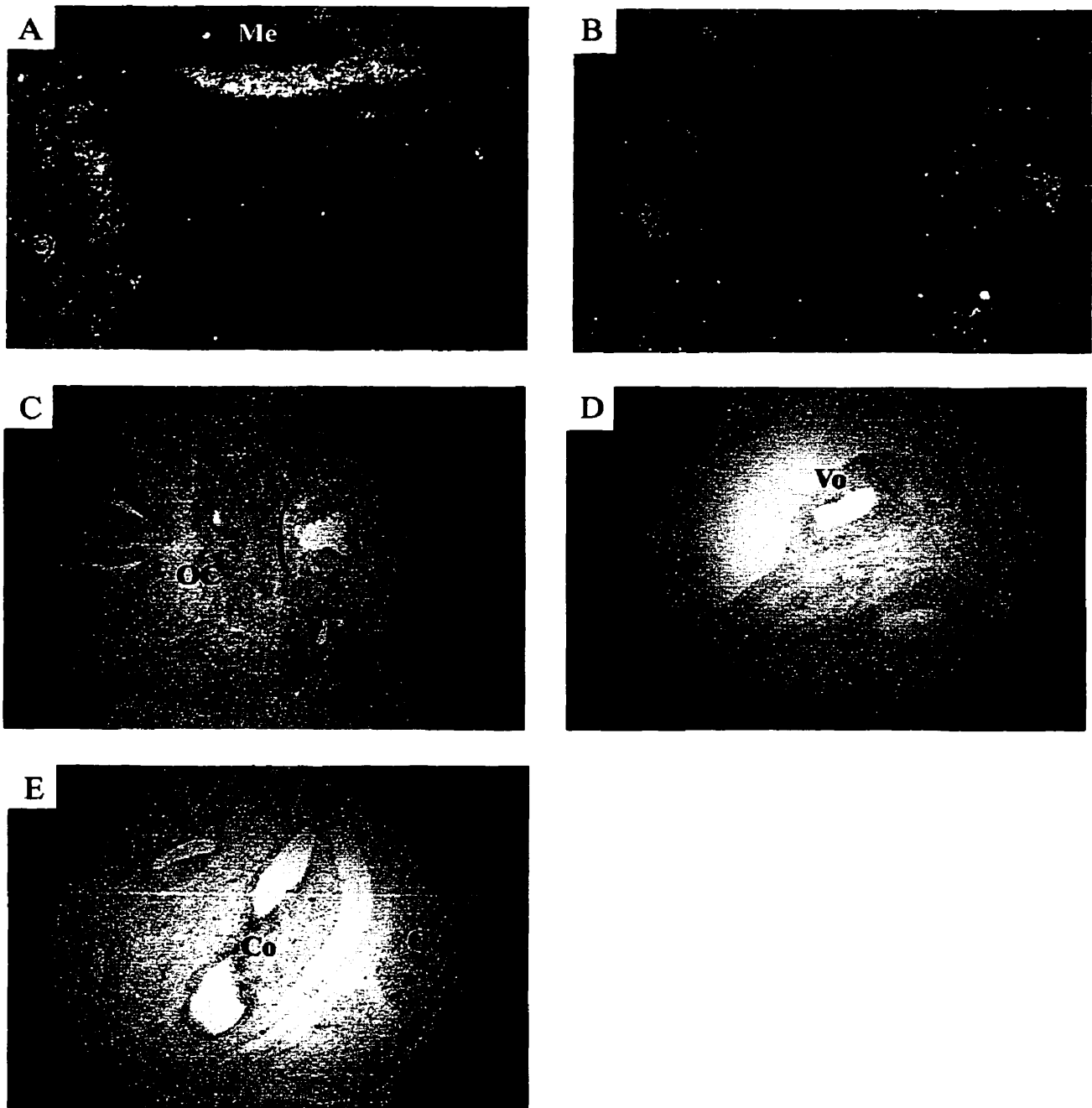
### **Chromosome mapping of *RTL-1* in human and mouse**

*RTL-1* was mapped to 18q22-23 on the human genome in collaboration with Dr. S.W. Scherer by radiation hybridization mapping (data not shown). This corresponds to



**Figure 2-11.** *Rtl-1* expression in the developing mouse spinal cord.

A) Expression of *Rtl-1* was first examined in the developing spinal cord (cervical level) and dorsal root ganglia at E13.5. Two strong areas of staining are present near the ventral horns. At E15.5 (B), *Rtl-1* expression is uniform throughout the spinal cord (cervical level). At E13.5, towards the thoracic region of the spinal cord, *Rtl-1* appears to be more strongly expressed in the ventral spinal cord (C, darkfield; D, brightfield), but is uniform at E15.5 (E). Dorsal root ganglia (DRG), dorsal (Dor).



**Figure 2-12.** Expression of *Rtl-1* in various neural regions during mouse embryogenesis.

*Rtl-1* expression can be detected in the future medulla oblongata at E13.5 (A) and in a small area of the presumptive optic chiasm at E12.5 (B). E15.5 (C), *Rtl-1* is expressed throughout the optic chiasm, trigeminal (V) ganglion, facial ganglion, olfactory epithelium, vomernasal organ (D), glossopharyngeal (IX) ganglia and cochlea (E). Vomernasal organ, (Vo); olfactory epithelium, (OE); cochlea, (Co); glossopharyngeal (IX) ganglia, (GG); future medulla oblongata, (Me); optic chiasm, (OC); trigeminal (V) ganglion, (TG); facial (VII) ganglion, (FG). (A and B, darkfield; C-E, brightfield illumination).

syntenic regions on chromosomes 1 and 18 in mouse. However, no genetic disorders involving the eye or nervous system map to this region.

## **DISCUSSION**

The Human Genome Project has generated a vast amount of data through its EST sequencing project. These sequences have been organized into overlapping clusters and can be sorted according to their pattern of tissue expression. In this work, an *in silico* screen was conducted to identify novel cDNAs which were likely to be important in retinal development.

After examination of 203 retina-only clusters, twenty-five clusters were known genes, of which, eight are important in eye development (seven of these are photoreceptor-specific). Only one cluster (Hs.60563; *RTL-1*) satisfied the original criteria of being a novel, retina-only cluster with homology to known developmental proteins. The majority of the clusters examined (170 of 204) were classified as novel because they did not have any similarity with known polypeptides. This result can be largely attributed to the fact that most of the sequences were from 3'UTR of the respective gene, which is generally not conserved between different species. This fact reduced the efficiency of the screen. Based on the expression pattern of the known genes, and despite the lack of coding sequences in most of the clusters, as many as 28% (7 of the 25 known genes are photoreceptor-specific) of the clusters may represent novel photoreceptor-specific genes (although this is a relatively small sample size). In a future *in silico* screen, with the recent completion of the rough human genome sequence, it will be possible to re-examine the novel retina-only UniGene sequences with human genomic sequences. Using exon

prediction algorithms, such as GRAIL (Genome Informatics Corporation) to generate hypothetical ORFs, it will be easier to determine whether these hypothetical proteins have any homology to known developmental proteins.

I identified a novel murine CUB domain-encoding gene, *Rtl-1*, using the *in silico* screen. Based on cDNA sequences and predicted splice acceptor/donor site usage, *Rtl-1* is likely to yield at least three different mRNA transcripts resulting in three different isoforms, Rtl-1A, -1B, and -1C. In the adult mouse, three different transcripts can be detected in neural tissues by northern blot analysis. These transcripts are all longer than the largest *Rtl-1* cDNA sequence identified and it remains to be determined which splice variant corresponds to which transcript on the northern blot. The identity of different transcripts can be determined by probing a northern blot with sequence specific to each of the *Rtl-1* splice variants.

Despite the large size discrepancies between the northern blot transcripts and the known *Rtl-1* cDNA sequences, it is likely that I have identified the entire coding sequences for Rtl-1A, -1B, and -1C because i) there is an in-frame stop codon located 12 bp upstream of the most 5' methionine codon, ii) no identifiable splice acceptor sites are present between the 5' stop codon and the start codon, iii) there is a predicted cleavable signal sequence at the amino terminus of the expected protein product, and iv) the nucleotide sequence between *Rtl-1* and *RTL-2*, which occurs only in the coding region, is highly conserved. Thus, the difference in size between the *Rtl-1* cDNAs and the transcripts detected on the northern blot are more likely to be due to missing 3'UTR and 5' UTR sequences. Incomplete 3'UTR sequence is most probable since none of the *Rtl-1* cDNAs has a fully conserved poly-adenylation consensus sequence ~20 bp from the 3'

poly(A) tail. These cDNAs may have arisen from mispriming of the oligo (d)T primer within an A-rich stretch in the 3'UTR. The 5' end of the *Rtl-1* cDNA sequences is GC-rich (68 % GC, 260/381 bp), suggesting that the region may be a CpG island, which is often found at the 5' region of genes (Lewin, 2000). However, it remains to be determined how much 5'UTR sequence is missing. To identify the remaining 5'UTR and 3'UTR of *Rtl-1*, experiments such as primer extension studies to determine the transcriptional start site(s), or 5' RACE and 3' RACE can be used to identify additional UTR sequence. These experiments will help to explain the size differences between the *Rtl-1* transcripts and their respective cDNA sequences.

The three *Rtl-1* isoforms all share a predicted secretory signal sequence and a CUB domain (CUB1), but differ at their carboxy termini. *Rtl-1A* has two CUB domains and an EGF-like motif, while *Rtl-1B* is a single CUB domain. Both lack a transmembrane region or a sequence for retention in the endoplasmic reticulum (Sanders and Schekman, 1992, Nielsen *et al.*, 1997) and thus are expected to be secreted from the cell. *Rtl-1C*, however, resembles *Rtl-1A* except that it contains a potential membrane spanning region and cytoplasmic domain. Thus, *Rtl-1C* is expected to anchor at the plasma membrane with its N-terminus exposed at the cell surface.

Amino acid sequence analysis shows that the CUB domains of *Rtl-1* are most homologous to the CUB domains in tolloids and neuropilins. Moreover, the domain organization of *Rtl-1A* and *Rtl-1C* resembles a hybrid between tolloids and neuropilins. The extracellular portion of both isoforms consists of two CUB domains and an EGF-like motif, much like the carboxy-terminus of tolloids (and *Bmp-1*), although *Rtl-1A* and *-1C* lack a metalloprotease domain. The domain organization of *Rtl-1C* also resembles the

neuropilins. Both proteins have two N-terminal CUB domains, a transmembrane region, and a cytoplasmic segment. Rtl-1C, however, lacks the two coagulation factor-like and MAM domains found in neuropilins. Additionally, neuropilins do not have an EGF-like motif (Figure 1-5). The similarity of the domain organization of Rtl-1 with tolloids and neuropilins suggests a function that may be related, at least partially, to each of those two classes of proteins. However, based on the expression pattern, the presence of a putative transmembrane domain, and the lack of a metalloprotease domain, Rtl-1C more closely resembles neuropilins than tolloids.

One potential function of Rtl-1C may be as a receptor for an as yet unknown ligand, perhaps in a signal transduction pathway. The two CUB domains at the amino-terminus are excellent candidates for ligand binding, as CUB domains found in other proteins are directly involved in protein-protein or protein-carbohydrate interactions. The cytoplasmic tail in Rtl-1C contains several serine and threonine residues, with putative phosphorylation sites for protein kinase C, cAMP- and cGMP-dependant protein kinases, and casein kinase II, which may serve as a mechanism of regulation. In addition, the last three residues T-R-V at the C-terminus of Rtl-1C comply with the PDZ-domain binding consensus sequence (S/T)-X-V-COOH. PDZ domains are often found in adapter proteins that mediate the formation of protein complexes. In several receptor signaling pathways, PDZ-containing proteins have been shown to interact directly with cell surface receptors, including ERB2/HER2 (Borg, 2000), EphB2 (Cowan, 2000), and neuropilin-1 (Cai and Reed, 1999). Therefore, these features suggest that Rtl-1C may resemble a cell surface receptor.



If Rtl-1A and the extracellular domains of Rtl-1C are involved in ligand binding, they may bind to a common ligand because they share the same extracellular domains. In such a scenario, the binding of secreted Rtl-1A to the hypothetical ligand(s) may impair the function of Rtl-1C as a potential receptor, as the Rtl-1C binding site on the ligand(s) may be sequestered.

Recently, a soluble form of neuropilin-1 was identified *in vivo* that consisted of the two CUB domains and coagulation factor-like domains. This isoform was shown to bind to VEGF<sub>165</sub> and antagonize VEGF<sub>165</sub> signaling (Gagnon *et. al.*, 2000) (possibly by interfering with VEGF<sub>165</sub> interaction with full-length neuropilin-1). The binding of this soluble isoform of neuropilin-1 to Sema3A, 3D, or 3E, has not yet been tested.

If such an antagonistic mechanism occurs between Rtl-1A and Rtl-1C, one potential method of controlling the amount of Rtl-1A and Rtl-1C protein may be through the regulation of their mRNA levels. In the adult retina for example, *Rtl-1C* transcripts predominate, as compared with *Rtl-1A* and *Rtl-1B* transcripts (Figure 2-6A and B), suggesting that Rtl-1C protein may be produced in higher abundance than the other isoforms (assuming that all isoforms are translated at the same rate and have the same half-life). Since a secreted protein occupies a larger volume (3D) than a membrane-bound protein (2D), the transcript levels of *Rtl-1A* observed in the adult retina may represent a low concentration of Rtl-1A protein in the extracellular environment, as the protein would probably diffuse and be diluted. It would be interesting to determine the relative levels of the different *Rtl-1* transcripts during embryonic development (which can be done by RT-PCR, although this will only look at a global embryonic picture).

*Rtl-1B* encodes only a single CUB domain (corresponding to CUB1 in *Rtl-1A* and *Rtl-1C*), a structure reminiscent of the spermadhesins, a family of proteins composed entirely of a signal sequence and a single CUB domain. Although spermadhesins are involved in protein-carbohydrate interactions, the sequence conservation between spermadhesins and *Rtl-1B* is low (~22% identity). The CUB domain of *Rtl-1B* has better conservation with the CUB domains that are involved in protein-protein interactions (~40% with tolloids and neuropilins), and thus *Rtl-1B* is more likely to be involved in a protein-protein complex rather than formation of a protein-carbohydrate complex. In addition, CUB domains have also been shown to form dimers, and *Rtl-1B* may also be involved in complex with other CUB domain-containing proteins. If such dimers form, the binding specificities conferred are likely to vary.

As a first step towards understanding the role of *Rtl-1* during development, the expression pattern in the developing mouse embryo was examined. Using the developing eye as a model for neural development, *Rtl-1* expression was first detected at E13.5 in post mitotic neurons - the committed and partially differentiated ganglion cells. Beginning at E14.5, and during E15.5, *Rtl-1* expression was detected in the neuroblastic layer, perhaps in the newly differentiated retinal neurons. The timing of expression in the presumptive GCL suggests that *Rtl-1* is turned on in post-mitotic neurons, and that later expression in the neuroblastic layer may be due to expression of other retinal cell types exiting their mitotic cycle (*i.e.* have become post-mitotic) but have not yet differentiated or migrated towards their respective nuclear layer. Alternatively, *Rtl-1* expression in the neuroblastic layer may be in neuroprogenitor cells. The use of double labeled *in situ* hybridizations using markers specific for retinal neuroprogenitors or differentiating

retinal neurons would help to clarify the identity of the cells expressing *Rtl-1* in the neuroblastic layer. In the mature retina, *Rtl-1* is expressed exclusively in the INL, where bipolar cells reside. However, no expression was observed in the GCL. The fact that *Rtl-1* is expressed in the mature retina suggests that this protein may also have a maintenance role in only some post-mitotic neurons.

*Rtl-1* is expressed along the entire length of the developing spinal cord during E13.5 and E15.5. This expression pattern suggests that *Rtl-1* may be important in the development of spinal cord neurons although additional studies will be required to define its role. Interestingly, at the cervical level of the spinal cord at E13.5, *Rtl-1* expression is uniform throughout the developing spinal cord along the dorsoventral axis with the exception of two regions of strong expression located near the ventral horns. Analysis using cell-specific markers could be used to clarify the identity of these cells. The thoracic/lumbar spinal cord, which is not as further developed as the cervical spinal cord, does not have the two strong bodies of *Rtl-1* expression. Instead, a gradient of *Rtl-1* expression can be detected, which is not evident in E13.5 spinal cord sections at the cervical level. At E14.5 (data not shown) and at E15.5, neither the strong *Rtl-1* staining at the cervical spinal cord or the gradient of expression in the thoracic/lumbar spinal cord is evident by *in situ* hybridization. It would be interesting to examine the cervical level of the spinal cord in earlier embryonic stages (particularly between E9.5 and E11.5, when spinal cord neurons are born and begin to migrate and extend axons) to determine whether the two regions of expression are only detected in a small window of development (at E13.5). The identity of these cells by future marker analysis may provide further clues into the possible role of *Rtl-1* at this stage of spinal cord

development. In addition, the examination of Rtl-1 protein by immunohistochemistry using isoform-specific anti-Rtl-1 antibodies at E13.5 can determine whether Rtl-1 exists in a gradient in the developing spinal cord.

*Rtl-1* is expressed strongly in various ganglia (a class of neurons which extend their axons over a relatively long distance) during mouse neurogenesis, including the dorsal root ganglia (DRG), trigeminal, glossopharyngeal, facial, and initially, retinal ganglion cells, suggesting that *Rtl-1* may be important for the formation of this class of neurons. Interestingly, at E12.5, and at E15.5, *Rtl-1* is expressed in the developing hypothalamus and the region of the optic chiasm. At E12.5, RGC growth cones exit the optic nerve and enter the hypothalamus (a part of the diencephalon). The axons then extend towards the ventral midline of the diencephalon (consisting of a population of radial glial cells), where axons must crossover the midline and form the optic tracts along the contralateral side, or steer away from the midline and project along the ipsilateral side of the hypothalamus (Marcus *et al.*, 1995).

Work in drosophila has identified several molecules including Roundabout (Robo, a guidance receptor) (Seeger *et al.*, 1993) and Slit (an inhibitory extracellular matrix protein secreted by midline glial cells) (Battye *et al.*, 1999) which play a role in axon guidance through the midline. In mammals, two Robo and three Slit homologs have been identified. To examine whether these molecules may be involved in RGC axon guidance, Erskine *et al.* (2000) examined the expression patterns of Robo1/2 and Slit1/2 in the developing mouse visual system. At E12.5, Robo2 is expressed in developing RGCs, while in the hypothalamus, where the RGC axons will migrate through, Robo1/2, and

Slit1/2 expression is present. These expression patterns are similar at E15.5, at a time when RGC axon crossover at the optic chiasm is highest (Erskine *et al.*, 2000).

Although the expression pattern of *Rtl-1* in the region of the developing optic chiasm does not directly overlap with the Robos and Slits (to confirm this, tissue sections through additional planes will need to be examined for *Rtl-1* expression), its spatial-temporal pattern of expression in this region is remarkably similar.

In the adult brain, *Rtl-1* is expressed in the olfactory apparatus and dentate gyrus of the hippocampus. These are areas of the mature nervous system where neurons are constantly replaced throughout adulthood (Linnarsson *et al.*, 2000), and are also regions of neuropilin expression (Chen *et al.*, 2000). Thus, the pattern of *Rtl-1* expression in the developing optic chiasm and in parts of the nervous system where neurons are renewed throughout adulthood, and its domain topological resemblance with the neuropilins, point to a speculative role of *Rtl-1* in axon projection or guidance. *Rtl-1* expression in the mature nervous system also suggests a role in neural maintenance.

Given that *Rtl-1* is expressed widely in the nervous system, it is interesting that *Rtl-1* is not expressed in the cerebellum. It is possible that *Rtl-1* may not have a role in the development or maintenance of cerebellar neurons, or perhaps, another molecule expressed in the cerebellum may have a function similar to that of *Rtl-1*.

A *Rtl-1* homolog with the same domain organization as *Rtl-1C* was identified and is referred to as *Rtl-2*. Preliminary expression analysis during mouse embryogenesis at E15.5 shows that *Rtl-2* has overlapping areas of expression with *Rtl-1*, including the cochlea, glossopharyngeal ganglia, and spinal cord. In the adult mouse, one region of expression that is unique to *Rtl-2* is the cerebellum. Due to the degree of sequence

conservation between Rtl-1C and Rtl-2, Rtl-2 may be able to compensate for the apparent lack of Rtl-1 in the cerebellum. Alternatively, Rtl-2 may have a distinct function to that of Rtl-1. The cerebellum would therefore be an ideal model to study the role of Rtl-2 in the absence of Rtl-1.

The presence of a *Rtl-1* homolog with overlapping expression is analogous to the expression of the neuropilins. Neuropilin-1 and neuropilin-2 are also coexpressed in various neuronal tissues and can form heterodimers. By analogy, *Rtl-1* and *Rtl-2* may also interact to form complexes that may lead to different binding specificities with putative ligand(s). Dimerization of neuropilins is mediated by the MAM domain, which is not found in Rtl-1 or Rtl-2. If Rtl-1 and Rtl-2 form homo/heterodimers, the interaction may occur through a different mechanism - perhaps through different regions of Rtl-1 and Rtl-2 (*e.g.* the CUB domain or cytoplasmic domains), by an adaptor protein, or protein complex.

With the completion of the genomic sequence for several model organisms, it is now possible to quickly determine whether homologs of Rtl-1 exist. Predicted proteins with the same domain organization as Rtl-1C have been identified in *C. elegans* (K05C4.11) and *Drosophila* (CG5449). A hypothetical protein resembling Rtl-1A in *C. elegans*, consisting of a putative secretory signal sequence and two CUB domains (K03E5.1), was also identified. In addition, evidence from human genomic DNA suggests the presence of another Rtl-1 homolog, designated RTL-3.

From these studies, I have identified a novel family of putative cell membrane molecules that are conserved between invertebrates and vertebrates. While my initial studies in mouse suggest a role of Rtl-1 in neural development, and perhaps a speculative

role in axon guidance, the function of Rtl-1 and its family members are yet to be determined.

## **CHAPTER 3**

Future and Proposed Ph.D. Studies



## **FUTURE DIRECTIONS**

During my studies, I conducted an *in silico* screen that led to the cloning of a novel murine CUB domain-encoding gene, *Rtl-1*. As a first step towards understanding the biological role of *Rtl-1*, I examined its developmental and adult expression pattern in mice. While these studies have provided some clues into the potential biological role of *Rtl-1*, additional experiments are required to clarify some of these issues. These include the identification interacting partners, determination of the cellular localization of the different *Rtl-1* isoforms, and biochemical analysis.

For my Ph.D. studies, I plan to study the genetics of *Rtl-1* and *Rtl-2* in mice. I will generate and study *Rtl-1* and *Rtl-2* deficient mice, and explore the developmental expression pattern of *Rtl-1* and *Rtl-2* in greater detail. *Rtl-1*<sup>-/-</sup> *Rtl-2*<sup>-/-</sup> double mutants will also be examined. While the gene targeted mice are generated, mutants in *C. elegans* *Rtl-1* homologs will be isolated and preliminary studies on these animals will be initiated.

### **Identification of Rtl-1 interacting partners**

The identification of *Rtl-1* interacting partners will help to elucidate the molecular pathway(s) in which *Rtl-1* participates. CUB domains and EGF-like motifs are often involved in protein-protein interactions and thus are good starting points for *Rtl-1* interaction studies. To identify such interactors, the two *Rtl-1* CUB domains and EGF-like motif, fused in-frame with a GST tag and a phosphorylation site, will be used to screen a mouse brain phage expression library. The interaction between *Rtl-1* and

positive clones will be confirmed using other methods, such as affinity chromatography, or co-immunoprecipitation.

In addition to the extracellular portion of Rtl-1, the putative intracellular tail found in Rtl-1C is also a potential site of protein-protein interaction. To identify such interacting proteins, the putative cytoplasmic tail will be used to screen mouse developing eye E14.5 and adult retinal yeast two-hybrid libraries. The interaction of positive clones with Rtl-1 will be confirmed as in the phage expression library screen.

In conjunction with unbiased library screens, candidate molecules can also be tested for interaction with Rtl-1. Based on the similarity of the CUB domains and expression pattern of Rtl-1 with neuropilins, Rtl-1 may bind to a semaphorin-like molecule. Neuropilin-1 and -2 bind to class 3 semaphorins (Sema3A - F), and are expressed in retinal ganglion cells. An additional semaphorin, Sema4F, also repels RGCs *in vitro*. However, its receptor is unknown. While purely speculative, the CUB domains of Rtl-1C may interact with the extracellular domains of Sema4F, a transmembrane protein, and this interaction can be tested in tissue culture using a soluble Sema4F-AP (alkaline phosphatase) fusion protein to determine whether the recombinant protein stains the surface of cells transfected with Rtl-1C.

The putative cytoplasmic tail of Rtl-1C has a consensus PDZ binding site. The PDZ domain participates in protein-protein interaction and is found in some proteins that mediate the formation of protein-complexes (*i.e.* as an adapter protein). Therefore, PDZ proteins expressed in regions overlapping with *Rtl-1* expression are good interacting candidates. One possible interactor is neuropilin-1 interacting protein (NIP-1, also referred to as GIPC/SemaCAP1), a PDZ-containing protein expressed in overlapping

regions with *Rtl-1* in the developing eye. The interaction between NIP-1 and the putative Rtl-1C cytoplasmic tail could be tested by co-immunoprecipitation or yeast two-hybrid experiments.

In light of the biochemical studies that have shown direct interaction between neuropilin-1 and -2, Rtl-1 (particularly Rtl-1C) and Rtl-2 may also interact together. The interaction between the two can be examined in transfected cell lines through cross-linking and western blotting experiments using antibodies specific to Rtl-1C and Rtl-2.

### **Rtl-1 protein localization studies**

Protein localization studies will complement the *Rtl-1* expression data obtained by *in situ* hybridization. To examine the distribution of Rtl-1, isoform-specific anti-Rtl-1 antibodies could be raised. Western blot analysis could be performed to determine if the level of each Rtl-1 isoform changes at different stages of development or between different tissues.

By immunohistochemistry, the precise location of Rtl-1A/Rtl-1B (predicted to be secreted) and Rtl-1C (expected to localize on the cell surface) can be determined. The cellular localization of Rtl-1, particularly Rtl-1C, would provide a clue towards its function. For example, it would be interesting if Rtl-1C, like neuropilin-1 and -2, is localized to neuronal growth cones (a sensory protuberance located at the end of elongating axons, implying a possible axonal guidance function) or throughout the cell surface (which may suggest a role of Rtl-1 in a more global cellular context).

## **Post-translational modifications of Rtl-1**

Rtl-1 has a predicted secretory pathway signal sequence and phosphorylation sites. Glycosylation of the extracellular domains are also possible. To examine whether Rtl-1 has a cleavable signal sequence or is glycosylated, Rtl-1 will be translated *in vitro* using a rabbit reticulocyte lysate in the presence or absence of endoplasmic reticulum microsomal membranes. Samples of the translation products will be treated with endoglycosylase H (which cleaves high mannose structures from N-linked glycoproteins). Treated and untreated proteins will be analyzed by SDS PAGE and examined for changes in relative mobility characteristic of signal sequence cleavage and/or glycosylation.

To examine whether Rtl-1C is phosphorylated, *in vitro* transcribed and translated Rtl-1 protein could be incubated with  $\gamma^{32}\text{P}$ -labelled ATP and kinases that have predicted phosphorylation sites in Rtl-1. Rtl-1 phosphorylation can be assessed by SDS PAGE and autoradiography. The identification and mapping of phosphorylation sites maybe important for understanding Rtl-1C regulation *in vivo*, particularly if it is a membrane receptor.

## **PROPOSED Ph.D. STUDIES**

The overall goal of my Ph.D. studies will be to understand the role of *Rtl-1* and *Rtl-2* in the developing mouse nervous system using the developing neuroretina and spinal cord as a model. Currently, experiments to identify Rtl-1 interacting partners and to understand the biochemistry of Rtl-1 and Rtl-2 are being pursued by another member in our lab. I will be studying the genetics of *Rtl-1* and *Rtl-2*. To approach this question, I

will generate *Rtl-1* and *Rtl-2* deficient mice and examine their phenotype(s). In addition, I will perform a more in-depth analysis of the developmental expression pattern of both *Rtl-1* and *Rtl-2*, and isolate *C. elegans Rtl-1* homologs.

### **Specific Aims**

- 1) Using homologous recombination techniques, I will generate *Rtl-1*<sup>-/-</sup>, *Rtl-2*<sup>-/-</sup> mice, and *Rtl-1*<sup>-/-</sup> *Rtl-2*<sup>-/-</sup> double mutants (generated by crossing *Rtl-1*<sup>+/-</sup> *Rtl-2*<sup>+/-</sup> mutants).
- 2) To characterize the phenotype(s) of *Rtl-1*<sup>-/-</sup>, *Rtl-2*<sup>-/-</sup>, and *Rtl-1*<sup>-/-</sup>*Rtl-2*<sup>-/-</sup> mice, I will examine these mice for any morphological and/or histological abnormalities as a result of the introduced mutation(s).
- 3) By *in situ* hybridization and X-gal staining (using a *tau-lacZ* reporter gene), I will examine in greater detail, the developmental expression patterns of *Rtl-1* and *Rtl-2*.
- 4) To study the phenotype of the *C. elegans Rtl-1* homologs, mutant animals will be isolated and examined for any morphological and locomotive defects.

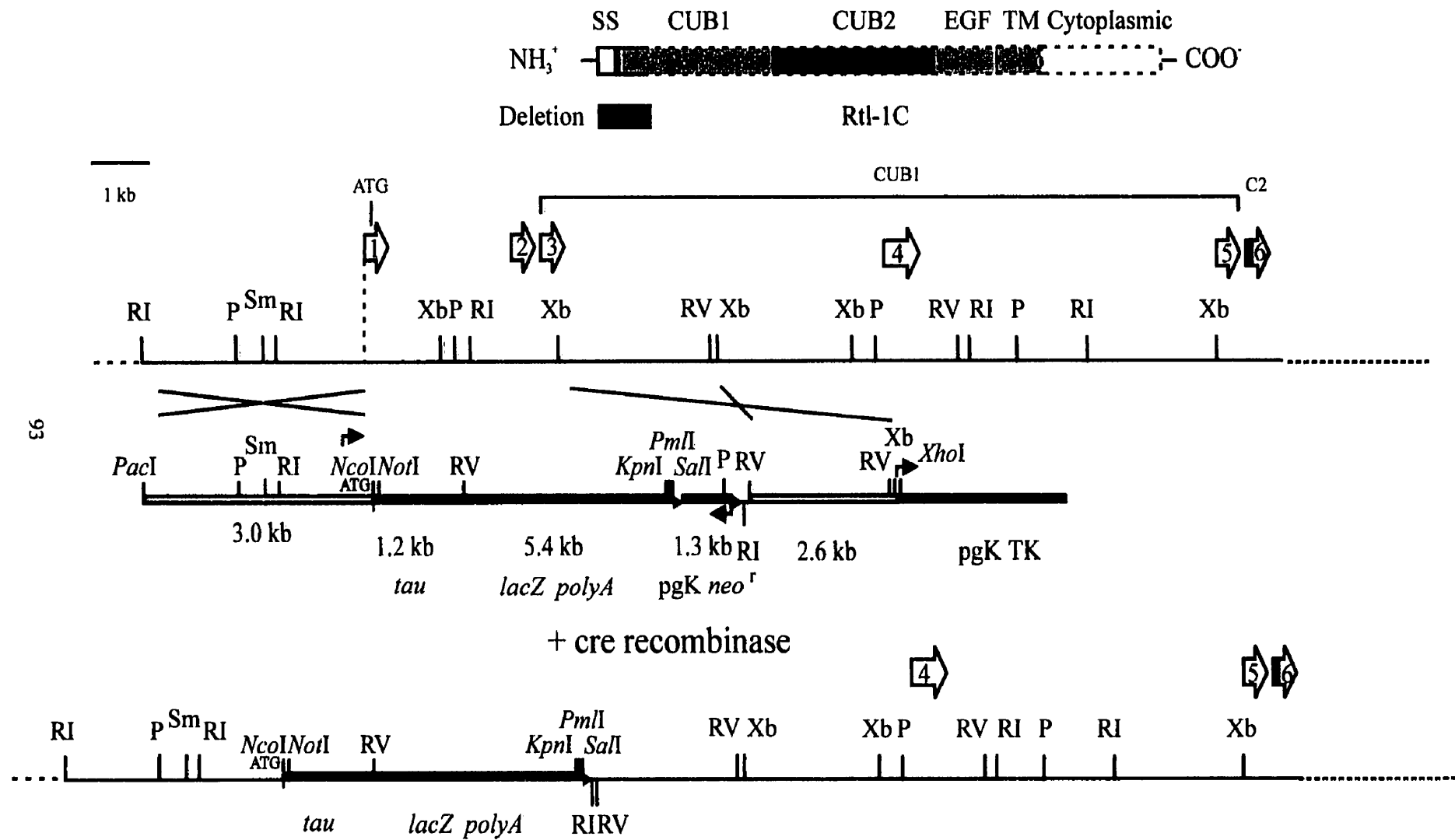
### 1.1) *Rtl-1* and *Rtl-2* gene targeting vectors

The mouse is an ideal model to study mammalian neural development, partly because of the relative ease of introducing gene-specific mutations. No naturally occurring phenotypes map to the *Rtl-1* or *Rtl-2* locus. Therefore, to begin to understand the function of *Rtl-1* and *Rtl-2* in neural development, mouse gene targeting experiments will be conducted.

Prior to construction of a *Rtl-1* gene targeting vector, the intron/exon boundaries of *Rtl-1* were determined. *Rtl-1* positive mouse (129/SvEvTac) genomic BACs and lambda clones were isolated. Fragments of these clones were subcloned and the intron/exon boundaries were determined by partial DNA sequencing and PCR.

A gene targeting vector designed to disrupt the 5' region of *Rtl-1* (deleting the predicted secretory signal sequence and first half of CUB1, thereby disrupting the *Rtl-1A*, *-1B*, and *-1C* isoforms) will be used. Since the expression pattern of *Rtl-1* is largely restricted to neurons, a *tau-lacZ* reporter gene will be introduced in-frame with the methionine start codon of *Rtl-1* (Figure 3-1). The *tau-lacZ* reporter, ideally, will label axon projections of *Rtl-1* expressing neurons, and provide a better understanding of the role of *Rtl-1* during murine embryogenesis.

Embryonic stem (ES) cells from a *RI* (129/SvCP) background will be electroporated with the linearized targeting vector. Cells surviving both positive (neomycin resistance) and negative (herpes simplex virus-1 thymidine kinase) selection will be screened by Southern blot hybridization for homologous recombination using flanking genomic probes. Clones will be confirmed by Southern blot analysis and PCR, and aggregation experiments will be performed by the Mouse Core Facility (The Hospital



**Figure 3-1.** Gene targeting strategy to disrupt *Rtl-1* by homologous recombination.

The linearized targeting vector will be electrophorated into RI ES cells, and positive clones will be identified by Southern blot and PCR analyses. Transcription from the *neo<sup>r</sup>* promoter is in the opposite direction of *Rtl-1* transcription. The cassette is flanked by loxP sites (solid black arrows) and will be removed by cre recombinase. The *Rtl-2* intron/exon boundaries are almost identical to *Rtl-1*, and therefore, a similar gene targeting strategy will be used to disrupt *Rtl-2*. Restriction enzyme sites: P, *Pst*I; Sm, *Sma*I; RI, *Eco*RI; RV, *Eco*RV; Xb, *Xba*I.

For Sick Children, Toronto, Canada) to generate chimeric mice. Standard breeding procedures will be performed to identify germline transmission, and these mice will be bred to obtain homozygous *Rtl-1* null mice. Since transcription of the neomycin resistance gene may interfere with transcription of the reporter gene, *Rtl-1* mutant mice, if necessary, will be bred with cre expressing mice to excise the neomycin gene, which are flanked by loxP sites.

To generate *Rtl-2* null mice, a similar strategy as described above will also be used. Like the *Rtl-1* construct, the *Rtl-2* targeting vector will incorporate an in-frame *tau-lacZ* reporter gene. The intron/exon boundaries of *Rtl-2* have essentially been determined, using human genomic *RTL-2* BAC sequences as a guide. (The gene structure between human *RTL-1* and mouse *Rtl-1* is largely conserved). Mouse 129/SvEvTac genomic BAC clones which contain the 5' end of *Rtl-2* are currently being examined by PCR and DNA sequencing to determine the size of the relevant introns.

Analogous with the *Rtl-1* targeting vector, the *Rtl-2* vector will be designed to disrupt the 5' portion of the gene and delete the anticipated signal sequence and first half of CUB1. Although it is not yet known whether *Rtl-2* alternative splice products are generated like *Rtl-1*, RT-PCR experiments will be performed to address this issue. If *Rtl-2* generates isoforms related to those as predicted for *Rtl-1*, the *Rtl-2* targeting vector should also disrupt these variants.

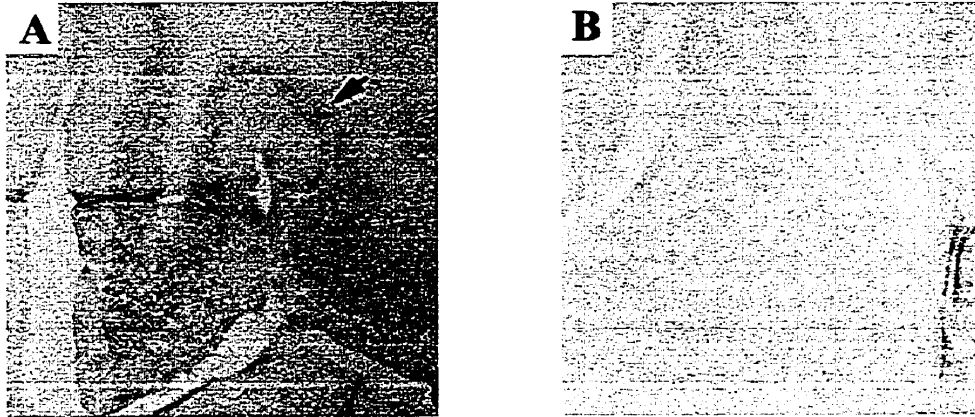
## **2.1) *Rtl-1*<sup>-/-</sup> and *Rtl-2*<sup>-/-</sup> phenotypic analysis**

*Rtl-1*<sup>-/-</sup> and *Rtl-2*<sup>-/-</sup> embryos and mice will be examined for developmental defects. This will include a morphological survey to look for any gross abnormalities and a



histological analysis of different parts of the nervous system. In addition, defects in neural development will be examined, and particular attention will be placed on developing neurons (*e.g.* to determine whether axon projections have been perturbed, or if the number of different neuronal cell types is affected). Tau-lacZ is expected to stain axon projections of *Rtl-1* expressing neurons. However, if the reporter is non-functional in the transgenic mouse (perhaps due to deletion of an enhancer element), whole mount embryos (up to ~E13.5) and embryonic sections will be stained with anti-neurofilament antibodies (such as 2H3) to visualize axons. The staining pattern will be compared with those of *Rtl-1*<sup>+/*tau-lacZ*</sup> and *Rtl-2*<sup>+/*tau-lacZ*</sup> mice, in addition to wild type *Rtl-1* and *Rtl-2* positive neurons (stained with 2H3). If *Rtl-1* and/or *Rtl-2* null mice have an axon projection phenotype, the 2H3 antibody will be useful to compare the axons of wild type and heterozygous mice to determine if there is any apparent dosage effect. To determine whether different spinal cord neuronal populations are affected in the mutants, cell-specific markers will be used to examine the cell numbers and their position in the spinal cord.

From preliminary studies, *Rtl-1* and *Rtl-2* are often co-expressed in similar tissues. In the cerebellum, however, only *Rtl-2* expression has been detected (Figure 3-2). Therefore, the cerebellum would be an ideal model for understanding the biological role of *Rtl-2* in the absence of *Rtl-1*. This tissue will be examined for morphological and histological aberrations in *Rtl-2* deficient mice.



**Figure 3-2.** *In situ* hybridization of adult mouse cerebellum sections.

(A) Section probed with *Rtl-2* antisense probe reveals purple staining in the cerebellar lobes as shown by the arrow, but no signal was observed using a *Rtl-1* antisense probe (not shown). No specific signal was detected with either *Rtl-1* (not shown) or *Rtl-2* (B) sense riboprobes.

## **2.2) Analysis of *Rtl-1*<sup>-/-</sup> *Rtl-2*<sup>-/-</sup> double mutants**

The high degree of amino acid identity and large regions of overlapping expression during development between *Rtl-1* and *Rtl-2* raises the possibility of the two proteins sharing similar, and perhaps in some cases, redundant function(s). If the two proteins share redundant functions, it is possible that *Rtl-1*<sup>-/-</sup> and *Rtl-2*<sup>-/-</sup> mutants may not produce any significant phenotype, and this would necessitate the generation of *Rtl-1*<sup>-/-</sup> *Rtl-2*<sup>-/-</sup> double mutants. These mice (double mutants), like the *Rtl-1* and *Rtl-2* single mutants, will be examined for defects in development or for any obvious behavioral changes. If double mutants are not viable, then embryos will be studied, as described above, to determine the cause of the embryonic lethal phenotype.

## **3.1) In-depth analysis of the *Rtl-1* and *Rtl-2* developmental expression pattern**

A preliminary study of the developmental expression pattern of *Rtl-1* is presented in this thesis. To further understand the role of *Rtl-1* and *Rtl-2* in development, more detailed expression analysis is required. Different staged wild type mouse embryos will be examined for *Rtl-1* and *Rtl-2* expression by *in situ* hybridization. In particular, important time points during retinal and neural tube neurogenesis will be examined. Cell-specific markers will also be used to confirm the identity of *Rtl-1* and *Rtl-2* expressing cells in the developing eye and neural tube. Thus, additional expression information may provide further insight towards which tissues to examine more closely during the mutant phenotypic analysis.

In conjunction with *in situ* hybridization experiments, expression data provided by the *tau-lacZ* reporter gene (if the reporter gene is functional and representative of *Rtl* gene

expression) will also be used. Heterozygous *Rtl-1*<sup>+/*tau-LacZ*</sup> and *Rtl-2*<sup>+/*tau-LacZ*</sup> mice will be stained with X-gal to visualize the pattern of *Rtl-1* expressing cells during development.

#### 4.1) Examination of *Rtl-1* homologs in other *C. elegans*

The genomic sequence for several model organisms, including *Drosophila* and *C. elegans*, is now complete. A search for molecules resembling Rtl-1C has identified homologs in *Drosophila* (CG5449) and *C. elegans* (K05C4.11 and K03E5.1). Although a *Drosophila* Rtl-1C homolog is present, no characterized mutants have been mapped near the CG5449 locus. In addition, mapped P elements are not within the vicinity to quickly disrupt the gene. In contrast, a mutagenized *C. elegans* library is available. Therefore, because of the availability of this reagent, I will focus my effort on identifying alleles of *Rtl-1* homologs in *C. elegans*. *C. elegans* is also an ideal model to study Rtl-1 homologs because the developmental fate of each cell in the animal is known, thus simplifying cell fate determination studies. In addition, a large number of mutants are available which can be used in genetic studies.

Two *C. elegans* Rtl-1 homologs are present (K05C4.11 and K03E5.1). The domain organization of K05C4.11 most closely resembles Rtl-1C, which consists of a predicted secretory signal sequence, two CUB domains, an EGF-like motif, TM domain, and a cytoplasmic tail. K03E5.1, has greater resemblance with Rtl-1A, and has a putative secretory signal sequence and two CUB domains. BLAST alignments show that K03E5.1 has significant amino acid identity with Rtl-1A (Figure 3-3).

To gain further insight into the potential function of *Rtl-1*, *C. elegans* mutants will be identified and characterized in collaboration with Dr. Joe Culotti (Mount Sinai

Hospital, Toronto, Canada). A *C. elegans* mutagenized library will be screened by using gene-specific primers to identify animals with a deletion in the K05C4.11 or K03E5.1 locus. If no mutant exists in the library, then a mutant could be obtained through the *C. elegans* Gene Knockout Consortium (BC Cancer Agency, Vancouver, Canada). The mutations will be verified by Southern blot and DNA sequencing, and homozygous mutants will be bred. The mutant will be examined for any obvious morphological malformations and analyzed for any motor defects. The phenotype(s) of this mutant might be useful in directing future studies of mammalian *Rtl-1* and *Rtl-2*.

```

Rtl-1A      M-----IYGRSLFHIIAS--LIILHSSGATKKGTEKQITPETQKSV-----QCGTWTk--HAEGGVETSPNYPSK   61
K03E5.1     MKNVCTSFYAQFFLTKIETRKQILDIMNGQEKKPTLRAILKLRNYVTHSTGSPPPQFDQCLSIKIGIGLQNFNIETSPNFPDR   84
              * .      * .      * .      * ** * . *      *      *      *      *      *      *      *      *
              . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .

Rtl-1A      YPPDRECVYIIEAAPRQCIELYFDEKYSIEPSW-----ECKFDHIEVRDGPFGFSPIIGRECGQNPP-VIKSSGREFLWIKF   137
K03E5.1     YPPNIDCVRVIHSRPQHDVVVKFHHVFHIESYDKIDAGEECPNDFIEFRDGRYGFSPLIARFCGDRMPKREIRAVSGFLWIRF   168
              *** .** . * . * . . . * . ** .      ** * ** *** .*** * ** * . * * . . ** * . *
              . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .

Rtl-1A      FADGELESMGFSARYNFTPDPDFKDLGVLKPLP[shaded] 221
K03E5.1     RSDSMLEYQGFSAEYAIVP----SKTGRFNNHE[shaded] 248
              . * ** *** * *      *      *      . . * . * . . . .      ** . ** * . * . *      . .
              . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .

Rtl-1A      [shaded] CEGNTFFCHSNM   303
K03E5.1     [shaded] CKLF-----   321
              . * * . *** ** * * . * * . ** . *** . . . * . . . . .      ** . * . *
              . .      . .      . .      . .      . .      . .      . .      . .      . .      . .      . .

Rtl-1A      CINNTLVCNGLQNCVYPWDENHCKDQTAS   332
K03E5.1     -----

```

100

**Figure 3-3.** Amino acid alignment between Rtl-1A and K03E5.1, a *C. elegans* predicted ORF.

K03E5.1 has 31% identity and 50% similarity with Rtl-1A. (The EGF-like motif was excluded from the calculation because it is not present in the K03E5.1 ORF). Identical amino acids are indicated with an asterisk below, and similar residues are marked with a period. Gaps are represented by a hyphen. CUB1 and CUB2 sequences are shaded yellow and green, respectively. The conserved cysteine residues in the CUB domains are shown in purple.

## REFERENCES

- Aert, R., Volckaert, G., McDougall, R.C., Rajandream, M.A. and Barrell, B.G. (2000). Conserved hypothetical protein (*Schizosaccharomyces pombe*). (Unpublished).
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Alvarez-Bolado, G., Schwarz, M., and Gruss, P. (1997). Pax-2 in the chiasm. *Cell Tissue Res.* 290:197-200.
- Anderson, R.E., Fisher, S.K., and Steinberg, R.H. (1978). Mammalian cones: disk shedding, phagocytosis, and renewal. *Investigative Ophthalmology and Visual Science* 17:117-133.
- Asano-Miyoshi, M., Kusakabe, Y., Abe, K., Emori, Y. (1998). Identification of taste-tissue-specific cDNA clones from a subtraction cDNA library of rat circumvallate and foliate papillae. *J. Biochem (Tokyo)*. 124(5):927-33.
- Auffray, C., Ansorge, W., Ballabio, A., Estivill, X., Gibson, K., Lehrach, H., Poustka, A. and Lundeberg, J. (1999). The European IMAGE consortium for integrated Molecular analysis of human gene transcripts. (Unpublished).
- Bagnard, D., Lohrum, M., Uziel, D., Puschel, A.W., and Bolz, J. (1998). Semaphorins act as attractive and repulsive guidance signals during the development of cortical projections. *Development* 125:5043-5053.
- Battye, R., Stevens, A., and Jacobs, J.R. (1999). Axon repulsion from the midline of the *Drosophila* CNS requires slit function. *Development*. 126:2475-2481.
- Besharse, J.C. (1986). Photosensitive membrane turnover: Differentiated membrane domains and cell-cell interaction. *The retina, Part I* 297-352.
- Bisgrove, B.W., Andrews, M.E., Raff, R.A. (1991). Fibropellins, products of an EGF repeat-containing gene, form a unique extracellular matrix structure that surrounds the sea urchin embryo. *Dev Biol.* 146(1):89-99.
- Boguski, M.S., Lowe, T.M.J., and Tolstoshev, C.M. (1993). dbEST - database for "expressed sequence tags". *Nat. Genet.* 4(4):332-3.
- Boguski, M.S., Tolstoshev, and C.M., Bassett, D.E. Jr. (1994). Gene discovery in dbEST. *Science.* 265(5181):1993-4.
- Bonaldo, M.F., Lennon, G., and Soares, M.B. (1996). Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res.* 6(9):791-806.

- Bone-Larson, C., Basu, S., Radel, J.D., Liang, M., Perozek, T., Kapousta-Bruneau, N., Green, D.G., Burmeister, M., and Hankin, M.H. (2000). Partial rescue of the ocular retardation phenotype by genetic modifiers. *J Neurobiol.* 42(2):232-47.
- Borg, J.P., Marchetto, S., Le Bivic, A., Ollendorff, V., Jaulin-Bastard, F., Saito, H., Fournier, E., Adelaide, J., Margolis, B., and Birnbaum, D. (2000). ERBIN: a basolateral PDZ protein that interacts with the mammalian ERBB2/HER2 receptor. *Nat. Cell Biol.* 2(7):407-414.
- Bork, P., and Beckmann, G. (1993). The CUB domain. *J. Mol. Biol.* 231:539-545.
- Briscoe, J., Pierani, A., Jessell, T.M., and Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell.* 101:435-445.
- Briscoe, J., Sussel, L., Serup, P., Hartigan-O'Connor, J., Jessell, T.M., Rubenstein, J.L., and Ericson, J. (1999). Homeobox gene Nkx2.2 and specification of neuronal identity by graded Sonic hedgehog signaling. *Nature.* 398:622-627.
- Bronner-Fraser, M., and Fraser, S. (1989). Developmental potential of avian trunk neural crest cells in situ. *Neuron.* 3:755-766.
- Bugert, P., Hanke, S., Chudek, J. and Kovacs, G. (2000). Analysis of a putative tumor suppressor gene region of 100 kb at chromosome 3p21.1 in conventional renal cell carcinoma. (Unpublished).
- Burmeister, M., Novak, J., Liang, M.Y., Basu, S., Ploder, L., Hawes, N.L., Vidgen, D., Hoover, F., Goldman, D., Kalnins, V.I., Roderick, T.H., Taylor, B.A., Hankin, M.H., and McInnes, R.R. (1996). Ocular retardation mouse caused by Chx10 homeobox null allele: impaired retinal progenitor proliferation and bipolar cell differentiation. *Nat. Genet.* 12:376-384.
- Cai, H., and Reed, R.R. (1999). Cloning and characterization of neuropilin-1 interacting protein: A PSD-95/Dlg/ZO-1 domain-containing protein that interacts with the cytoplasmic domain of neuropilin-1. *J. Neuro. Sci.* 19(15):6519-6527.
- Carter-Dawson, L.D., LaVail, M.M. (1979). Rods and cones in the mouse retina. I. Structural analysis using light and electron microscopy. *J. Comp. Neurol.* 188(2):245-62.
- Cepko, C.L., Austin, C.P., Yang, X., Alexiades, M., and Ezzeddine, D. (1996). Cell fate determination in the vertebrate retina. *Proc. Natl. Acad. Sci. USA* 93(2):589-95.



Chen, H., Bagri, A., Zupicich, J.A., Zou, Y., Stoeckli, E., Pleasure, S.J., Lowenstein, D.H., Skarnes, W.C., Chedotal, A., Skarnes, W.C., and Tessier-Lavigne M. (2000). Neuropilin-2 regulates the development of selective cranial and sensory nerves and hippocampal mossy fiber projections. *Neuron*. 25(1):43-56.

Chen, H., Chedotal, A., He, Z.G., Goodman, C.S., and Tessier-Lavigne, M. (1997). Neuropilin-2, a novel member of the neuropilin family, is a high affinity receptor for the semaphorins Sema E and Sema IV but not Sema III. *Neuron*. 19:547-559.

Clark, T.G., Conway, S.J., Scott, I.C., Labosky, P.A., Winnier, G., Bundy, J., Hogan, B.L., and Greenspan, D.S. (1999). The mammalian Tolloid-like 1 gene, *Tll1*, is necessary for normal septation and positioning of the heart. *Development*. 126(12):2631-42.

Cohen, A.I. (1970). Further studies on the question of the patency of saccules in outersegments of vertebrate photoreceptors. *Vision Research*. 10:445-453.

Cowan, C.A., Yokoyama, N., Bianchi, L.M., Henkemeyer, M., Fritsch, B. (2000). EphB2 guides axons at the midline and is necessary for normal vestibular function. *Neuron* 26(2):417-30.

Craft, C.M., Whitmore, D.H. and Wiechmann, A.F. (1994) Cone arrestin identified by targeting expression of a functional family. *J. Biol. Chem.* 269, 4613-4619.

Day, A.J. (1999). The structure and regulation of hyaluronan-binding proteins. *Biochem Soc Trans.* 27(2):115-21.

Delgadillo-Reynoso, M.G., Rollo, D.R., Hursh, D.A., and Raff, R.A. (1989). Structural analysis of the uEGF gene in the sea urchin *Strongylocentrotus purpuratus* reveals more similarity to vertebrate than to invertebrate genes with EGF-like repeats. *J Mol Evol* 29(4):314-27.

Devriendt, K., Van den Berghe, H., Cassiman, J.J. and Marynen, P. (1991). Primary structure of pregnancy zone protein. Molecular cloning of a full-length PZP cDNA clone by the polymerase chain reaction. *Biochim. Biophys. Acta* 1088 (1), 95-103

Dewji, N.N., Wenger, D.A. and O'Brien, J.S. (1987). Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 precursor. *Proc. Natl. Acad. Sci. U.S.A.* 84 (23), 8652-8656.

Dias, J.M., Carvalho, A.L., Kolln, I., Calvete, J.J., Topfer-Petersen, E., Varela, P.F., Romero, A., Urbanke, C., and Romao, M.J. Crystallization and preliminary X-ray diffraction studies of aSFP, a bovine seminal plasma protein with a single CUB domain architecture. *Protein Sci.* 6(3):725-7.

- Ding, L., Traer, E., McIntyre, T.M., Zimmerman, G.A. and Prescott, S.M. (1998). The cloning and characterization of a novel human diacylglycerol kinase, DGKiota. *J. Biol. Chem.* 273 (49), 32746-32752.
- Ding, Q., Motoyama, J., Gasca, S., Mo, R., Sasaki, H., Rossant, J., and Hui, C.C. (1998). Diminished Sonic hedgehog signaling and lack of floor plate differentiation in *Gli2* mutant mice. *Development.* 125:2533-2543.
- Dowling, J.E. *The retina. An approachable part of the brain.* Belknap press of Harvard University press, Cambridge, Massachusetts, 1987.
- Drager, U.C. (1985). Birth dates of retinal ganglion cells giving rise to the crossed and uncrossed optic projections in the mouse. *Proc. R. Soci. London* 224:57:77.
- Drager, U.C., and Olsen, J.F. (1980). Origin of crossed and uncrossed retinal projections in pigmented and albino mice. *J. Comp. Neurol.* 191:383-412.
- Duboule, D., ed. (1994). *Guidebook to the Homeobox Genes.* Oxford University Press, New York.
- Eichele, G. (1992). Budding thoughts. *The Sciences.* 30-36.
- Eisen, J.S. (1999). Patterning motoneurons in the vertebrate nervous system. *Trends neurosci.* 22(7):321-326.
- Ericson, J., Rashbass, P., Schedl, A., Brenner-Morton, S., Kawakami, A., van Heyningen, V., Jessell, T.M., and Briscoe, J. (1997). Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. *Cell.* 90:169-180.
- Eudy, J.D., Weston, M.D., Yao, S.F., Hoover, D.M., Rehm, H.L., Ahmad, I., Ma-Edmonds, M., Yan, D., Cheng, J.J., Beisel, K.W., Ayuso, C., Cremers, C., Davenport, S., Moller, C., Talmadge, C.B., Tamayo, M., Swaroop, A., Morton, C.C., Kimberling, W.J. and Sumegi, J. (1998). Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. *Science* 280 (5370), 1753-1757.
- Ewing, B., and Green, P. (2000). Analysis of expressed sequence tags indicates 35,000 human genes. *Nat. Genet.* 25:232-234.
- Ferda-Percin, E., Ploder, L.A., Yu, J.J., Arici, K., Horsford, D.J., Rutherford, A., Bapat, B., Cox, D.W., Duncan, A.M., Kalnins, V.I., Kocak-Altintas, A., Sowden, J.C., Traboulsi, E., Sarfarazi, M., and McInnes R.R. (2000). Human microphthalmia associated with mutations in the retinal homeobox gene CHX10. *Nat Genet.* 25(4):397-401.

- Fero, M.L., Rivkin, M., Tasch, M., Porter, P., Carow, C.E., Firpo, E., Polyak, K., Tsai, L.H., Broudy, V., Perlmutter, R.M., Kaushansky, K., and Roberts, J.M. (1996). A syndrome of multiorgan hyperplasia with features of gigantism, tumorigenesis, and female sterility in p27(Kip1)-deficient mice. *Cell*. 85:733-744.
- Fienburg, A.P., and Vogelstein, B.A. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 1983:132:6-13.
- Fujisawa, H., Takagi, S., and Hirata, T. (1995). Growth-associated expression of a membrane protein, neuropilin, in *Xenopus* optic nerve fibers. *Dev Neurosci.* 17(5-6):343-9.
- Finelli, A.L., Xie, T., Bossie, C.A., Blackman, R.K., Padgett, R.W. The tolkin gene is a tolloid/BMP-1 homologue that is essential for *Drosophila* development. *Genetics.* 141(1):271-81.
- Freund, C.L., Gregory-Evans, C.Y., Furukawa, T., Papaioannou, M., Looser, J., Ploder, L., Bellingham, J., Ng, D., Herbrick, J.-A.S., Duncan, A., Scherer, S.W., Tsui, L.-C., Loutradis-Anagnostou, A., Jacobson, S.G., Cepko, C.L., Bhattacharya, S.S. and McInnes, R.R. (1997). Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (CRX) essential for maintenance of the photoreceptor. *Cell* 91 (4), 543-553.
- Freund, C.F., Horsford, D.J., and McInnes, R.R. (1996) Transcription factor genes and the developing eye: a genetic perspective. *Hum. Mol. Genet.* 5:1471-1488.
- Fritsch, B., Nichols, D.H., Echelard, Y., and McMahon, A.P. (1995). Development of midbrain and anterior hindbrain ocular motoneurons in normal and Wnt-1 knockout mice. *J. Neurobiol.* 27:457-469.
- Gabor-Miklos, G.L., and Rubin, G.M. (1996). The role of the genome project in determining gene function: insights from model organisms. *Cell.* 86:521-529.
- Gagnon, M.L., Bielenberg, D.R., Gechtman, Z., Miao, H.Q., Takashima, S., Soker, S., and Klagsbrun, M. (2000). Identification of a natural soluble neuropilin-1 that binds vascular endothelial growth factor: *in vivo* expression and antitumor activity. *Proc. Nat. Acad. Sci.* 97:2573-2578.
- Gan, L., Wang, S., Huang, Z., and Klein, W.H. (1999). POU domain factor Brn-3b is essential for retinal ganglion cell differentiation and survival but not for initial cell fate specification. *Develop. Biol.* 210:469-480.
- Gan, L., Xiang, M., Zhou, L., Wagner, D., Klein, W.H., and Nathans, J. (1996). POU domain factor Brn-3b is required for the development of a large set of retinal ganglion cells. *Proc. Natl. Acad. Sci. USA* 93:3920-3925.

Geier, G., and Zwillig, R. (1998). Cloning and characterization of a cDNA coding for *Astacus* embryonic astacin, a member of the astacin family of metalloproteases from the crayfish *Astacus astacus*. *Eur. J. Biochem.* 253(3):796-803.

Ghabrial, A.S., Ray, R. and Schupbach, T. (1998). Bem46-like protein (*Drosophila melanogaster*). (Unpublished).

Giger, R.J., Cloutier, J.F., Sahay, A., Prinjha, R.K., Levengood, D.V., Moore, S.E., Pickering, S., Simmons, D., Rastan, S., Walsh, F.S., Kolodkin, A.L., Ginty, D.D., and Geppert, M. (2000). Neuropilin-2 is required in vivo for selective axon guidance responses to secreted semaphorins. *Neuron.* 25(1):29-41.

Giger, R.J., Urquhart, E.R., Gillespie, S.K.H., Levengood, D.V., Ginty, D.D., and Kolodkin, A.L. (1998). Neuropilin-2 is a receptor for semaphorin IV: Insight into the structural basis of receptor function and specificity. *Neuron* 21:1079-1092.

Giguere, V., Yang, N., Segui, P. and Evans, R.M. (1988). Identification of a new class of steroid hormone receptors. *Nature* 331 (6151), 91-94.

Gilbert, S.F. (1997). *Developmental biology*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Golden, J.A. and Chernoff G.F. (1993). Intermittent pattern of neural tube closure in two strains of mice. *Teratology.* 47(1):73-80.

Goulding, M.D., Lumsden, A., and Gruss, P. (1993). Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. *Development.* 117:1001-1016.

Grossberger, D. (1987). Minipreps of DNA from bacteriophage lambda. *Nucleic Acids Res.* 15(16):6737

Halder, G., Callaerts, P., and Gehring, W.J. (1995). Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. *Science* 267:1788-1792.

Hall, Z.W. (1992). *An introduction to molecular neurobiology*. Sinauer Associates, Inc. Sunderland, Massachusetts.

He, Z., and Tessier-Lavigne, M. (1997). Neuropilin is a receptor for the axonal chemorepellent semaphorin III. *Cell* 90:739-751.

Hishida, R., Ishihara, T., Kondo, K., Katsura, I. (1996). hch-1, a gene required for normal hatching and normal migration of a neuroblast in *C. elegans*, encodes a protein related to Tolloid and Bmp-1. *EMBO J.* 15(16):4111-22.

Hofmann, K., Bucher, P., Falquet, L., and Bairoch A. (1999). The PROSITE database, its status in 1999. *Nuc. Acids. Res.* 27:215-219.

Hunter, D.D., Murphy, M.D., Olsson, C.V., and Brunken, W.J. (1992). S-laminin expression in adult and developing retinæ: a potential cue for photoreceptor morphogenesis. *Neuron* 8:399-413.

Imondi, R., Wideman, C., and Kaprielian, Z. (2000). Complementary expression of transmembrane ephrins and their receptors in the mouse spinal cord: a possible role in constraining the orientation of longitudinally projecting axons. *Development.* 127:1397-1410.

Inazu, T. (2000). Gamm1: cloning and characterization. (Unpublished).

Ingham, P.W. (1998). Transducing Hedgehog: the store so far. *EMBO J.* 17:3505-3511.

Ishikawa, K., Nagase, T., Suyama, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1998). Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. *DNA Res.* 5 (3), 169-176.

Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T. (2000). NEDO human cDNA sequencing project. (Unpublished).

Kaufman, M.H. (1992). *The Atlas of Mouse Development*. Academic Press, Harcourt Brace Javanovich, Toronto, Canada.

Kawasaki, T., Kitsukawa, T., Bekku, Y., Matsuda, Y., Sanbo, M., Yagi, T., and Fujisawa, H. (1999). A Requirement for Neuropilin-1 in Embryonic Vessel Formation. *Development.* 126:4895-4902.

Kaneko, A. (1970). Physiological and morphological identification of horizontal, bipolar, and amacrine cells in goldfish retina. *J. Physiol.* 207:623-633.

Katagiri, C., Maeda, R., Yamashika, C., Mita, K., Sargent, T.D., Yasumasu, S. (1997). Molecular cloning of *Xenopus* hatching enzyme and its specific expression in hatching gland cells. *Int. J. Dev. Biol.* 41(1):19-25.

Kawakami, A., Kitsukawa, T., Takagi, S., and Fujisawa, H. (1996). Developmentally regulated expression of a cell surface protein, neuropilin, in the mouse nervous system. *J. Neurobiol.* 29:1-17.

Kelly, M.W., Turner, J.K., and Reh, T.A. (1994). Retinoic acid promotes differentiation of photoreceptors *in vitro*. *Development* 120:2091-2102.

Kikuno, R., Nagase, T., Ishikawa, K., Hirosawa, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1999). Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res.* 6, 197-205.

Kitsukawa, T., Shimizu, M., Sanbo, M., Hirata, T., Taniguchi, M., Bekku, Y., Yagi, T., and Fujisawa, H. (1997). Neuropilin-semaphorin III/D-mediated chemorepulsive signals play a crucial role in peripheral nerve projection in mice. *Neuron.* 19(5):995-1005.

Kitsukawa, T., Shimono, A., Kawakami, A., Kondoh, H., and Fujisawa, H. (1995). Overexpression of a membrane protein, neuropilin, in chimeric mice causes anomalies in the cardiovascular system, nervous system and limbs. *Development.* 121(12):4309-18.

Kolodkin, A.L., Levengood, D.V., Rowe, E.G., Tai, Y.T., Giger, R.J., and Ginty, D.D. (1997). Neuropilin is a semaphorin III receptor. *Cell* 90:753-762.

Kolodkin, A.L., and Ginty, D.D. (1997). Steering clear of semaphorins: Neuropilins sound the retreat. *Neuron* 19:1159-1162.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15(20):8125-48.

Kristiansen, M., Kozyraki, R., Jacobsen, C., Nexø, E., Verroust, P.J., Moestrup, S.K. (1999). Molecular dissection of the intrinsic factor-vitamin B12 receptor, cubilin, discloses regions important for membrane association and ligand binding. *J. Biol. Chem.* 274(29):20540-4.

Kumar, J., and Moses, K. (1997). Transcription factors in eye development: a gorgeous mosaic? *Genes and Develop.* 11:2023-2028.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.

Levine, E.M., Close, J., Fero, M., Ostrovsky, A., and Reh, T.A. (2000). p27(Kip1) regulates cell cycle withdrawal of late multipotent progenitor cells in the mammalian retina. *Dev Biol.* 219(2):299-314.

- Levine, E.M., Roelink, H., Turner, J., and Reh, T.A. (1997). Sonic hedgehog promotes rod photoreceptor differentiation in mammalian retinal cells *in vitro*. *J. Neurosci.* 17(16):6277-6288.
- Lewin, B. (2000). *Genes VII*. Oxford University Press. Toronto
- Leytus, S.P., Kurachi, K., Sakariassen, K.S., and Davie, E.W. (1986). Nucleotide sequence of the cDNA coding for human complement C1r. *Biochemistry* 25(17):4855-63.
- Lhomond, G., Ghiglione, C., Lepage, T., Gache, C. (1996). Structure of the gene encoding the sea urchin blastula protease 10 (BP10), a member of the astacin family of Zn<sup>2+</sup>-metalloproteases. *Eur J Biochem.* 238(3):744-51.
- Liang, F., Hold, I., Pertea, G., Karamycheva, S., Salzberg, S.L., and Quackenbush, J. (2000). Gene index analysis of the human genome estimates approximately 120,000 genes. *Nat. Genet.* 25:239-240.
- Liem, K., Tremmi, G., Roelink, H., and Jessell, T.M. (1995). Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. *Cell.* 82:969-979.
- Lindskog, M. and Blomberg, J. (1997). Spliced human endogenous retroviral HERV-H env transcripts in T-cell leukaemia cell lines and normal leukocytes: alternative splicing pattern of HERV-H transcripts. *J. Gen. Virol.* 78:2575-2585.
- Linnarsson, S., Willson, C.A., Ernfors, P. (2000). Cell death in regenerating populations of neurons in BDNF mutant mice. *Brain Res. Mol. Brain Res.* 75(1):61-9.
- Litingtung, Y., and Chiang, C. (2000). Specification of ventral neuron types is mediated by an antagonistic interaction between Shh and Gli3. *Nat. Neurosci.* 3(10):979-985.
- Liu, I.S., Chen, J.D., Ploder, L., Vidgen, D., van der Kooy, D., Kalnins, V.I., and McInnes, R.R. (1994). Developmental expression of a novel murine homeobox gene (Chx10): evidence for roles in determination of the neuroretina and inner nuclear layer. *Neuron* Aug;13(2):377-93.
- Liu, L.P., and Deber, C.M. (1999). Combining Hydrophobicity and Helicity: A Novel Approach to Membrane Protein Structure Prediction. *Bioorg. & Med. Chem.* 7, 1-7.
- Marcus, R.C., Blazeski, R., Godement, P., and Mason, C.A. (1995). Retinal axon divergence in the optic chiasm: uncrossed axons diverge from crossed axons within midline glial specialization. *J. Neurosci.* 15(5):3716-3729.

- Marques, G., Musacchio, M., Shimell, M.J., Wunnenberg-Stapleton, K., Cho, K.W., and O'Connor, M.B. (1997). Production of a DPP activity gradient in the early *Drosophila* embryo through the opposing actions of the SOG and TLD proteins. *Cell*. Oct 31;91(3):417-26.
- Matise, M.P., Epstein, D.J., Park, H.L., Platt, K.A., and Joyner, A.L. (1998). Gli2 is required for induction of floor plate and adjacent cells, but not most ventral neurons in the mouse central nervous system. *Development*. 125:2759-2770.
- Matsushita, F., Miyawaki, A., Mikoshiba, K. (2000). Vomeroglandin/CRP-Ductin is strongly expressed in the glands associated with the mouse vomeronasal organ: identification and characterization of mouse vomeroglandin. *Biochem Biophys Res Commun*. 268(2):275-81.
- McIlwain, J.T. (1996). *An introduction to the biology of vision*. Cambridge University Press, Cambridge, United Kingdom.
- Miller, N. (1999). hypothetical protein R13F6.4 - *Caenorhabditis elegans*. (Unpublished).
- Moran, L.A., Scrimgeour, K.G., Horton, H.E., Ochs, R.S., and Rawn, D.J. (1994). *Biochemistry*. Prentice Hall, London.
- Nagase, T., Ishikawa, K., Kikuno, R., Hirosawa, M., Nomura, N. and Ohara, O. (1999). Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res*. 6, 337-345.
- Nagase, T., Ishikawa, K., Nakajima, D., Ohira, M., Seki, N., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1997). Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. *DNA Res*. 4 (2), 141-150.
- Nakamura, F., Tanaka, M., Takahashi, T., Kalb, R.G., and Strittmatter, S.M. (1998). Neuropilin-1 extracellular domains mediate semaphorin D/III-induced growth cone collapse. *Neuron*. 21(5):1093-100.
- Nathans, J. and Hogness, D.S. (1984). Isolation and nucleotide sequence of the gene encoding human rhodopsin. *Proc. Natl. Acad. Sci. U.S.A.* 81 (15), 4851-4855.
- NCBI (National Center for Biotechnology Information) <http://www.ncbi.nlm.nih.gov>
- Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering*, 10, 1-6.



- Normes, H.O., and Carry, M. (1978). Neurogenesis in spinal cord of mouse: an autoradiographic analysis. *Brain Res.* 159:1-16.
- Obata, S., and Usukura, J. (1992). Morphogenesis of the photoreceptor outer segment during postnatal development in the mouse (BALB/c) retina. *Cell Tissue Res.* 269(1):39-48.
- Oliver, G., and Gruss, P. (1997). Current views on eye development. *Trends in Neurosci.* 20(9):415-421.
- Otting, G., Qian, Y.Q., Billeter, M., Miller, M., Affolter, M., Gehring, W.J., and Wuthrich, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* 9:3085-3092.
- Park, H.L., Bai, C., Platt, K.A., Matisse, M.P., Beeghly, A., Hui, C.C., Nakashima, M., Joyner, A.L. (2000). Mouse *Gli1* mutants are viable but have defects in SHH signaling in combination with a *Gli2* mutation. *Development* 127(8):1593-605.
- Pei, Y.F., and Rhodin, J.A.G. (1970). The prenatal development of the mouse eye. *Anat. Rec.* 168:105-126.
- Peng, Y., Song, H., Zhou, J., Huang, Q., Dai, M., Mao, Y., Yu, Y., Xu, X., Luo, B., Chen, J. and Hu, R. (2000). Human neuronal cell death related gene in neuron-7 (DN-7). (Unpublished).
- Piccolo, S., Agius, E., Lu, B., Goodman, S., Dale, L., and De Robertis, E.M. (1997) Cleavage of Chordin by Xolloid metalloprotease suggests a role for proteolytic processing in the regulation of Spemann organizer activity. *Cell.* 91(3):407-16.
- Pierce, E.A., Quinn, T., Meehan, T., McGee, T.L., Berson, E.L., and Dryja, T.P. (1999). Mutations in a gene encoding a new oxygen-regulated photoreceptor protein cause dominant retinitis pigmentosa. *Nat. Genet.* 22 (3), 248-254.
- Puschel, A.W. (1996). The Semaphorins: a family of axonal guidance molecules? *Euro. J. of Neurosci.* 8:1317-1321.
- Puschel, A.W., Adams, R.H., and Betz, H. (1996). The sensory innervation of the mouse spinal cord may be patterned by differential expression of and differential responsiveness to semaphorins. *Mol. and Cell. Neurosci.* 7:419-431.
- Rattner, A., Sun, H., and Nathans, J. (1999). Molecular genetics of human retinal disease. *Annu. Rev. Genet.* 33:89-131.

- Rhodes, S. and Huckle, E. (1999). Novel human mRNA from chromosome 1, which has similarities to BAT2 genes. (Unpublished).
- Romero, A., Ramao, M.J., Varela, P.F., Kolln, I., Dias, J.M., Carvalho, A.L., Sanz, L., Topfer-Paterson, E., and Calvete, J.J (1997). The crystal structures of two spermadhesins reveal the CUB domain fold. *Nat. Struct. Biol.* 4:783-788.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning: A laboratory manual. Second edition.* Cold Spring Harbour Laboratory Press, New York.
- Sanders, S.L., and Schekman, R. (1992). Polypeptide translocation across the endoplasmic reticulum membrane. *J. Biol. Chem.* 267(20):13791-1394.
- Schmidt, J., Francois, V., Bier, E., and Kimelman, D. (1995) *Drosophila* short gastrulation induces an ectopic axis in *Xenopus*: evidence for conserved mechanisms of dorsal-ventral patterning. *Development.* 121(12):4319-28.
- Schoenwolf, G.C., and Smith, J.L. (1990). Mechanisms of neurulation: traditional viewpoint and recent advances. *Development.* 109:243-270.
- Scott, I.C., Blitz, I.L., Pappano, W.N., Imamura, Y., Clark, T.G., Steiglitz, B.M., Thomas, C.L., Maas, S.A., Takahara, K., Cho, K.W.Y., and Greenspan, D.S. (1999). Mammalian BMP-1/Tolloid-Related metalloproteinases, including novel family member Mammalian Tolloid-Like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. *Dev. Biol.* 213:283-300.
- Seeger, M., Tear, G., Ferres-Marco, D., and Goodman, C.S. (1993). Mutations affecting growth cone guidance in *Drosophila*: genes necessary for guidance toward or away the midline. *Neuron* 10:409-426.
- Shimell, M.J., Ferguson, E.L., Childs, S.R., and O'Connor, M.B. (1991). The *Drosophila* dorso-ventral patterning gene *tolloid* is related to human bone morphogenetic protein 1. *Cell* 67:469-481.
- Shiwaku, H.O., Wakatsuki, S., Mori, Y., Fukushige, S. and Horii, A. (1997). Alternative splicing of GTBP in normal human tissues. *DNA Res.* 4 (5), 359-362.
- Shukunami, C. and Hiraki, Y. (1998). Expression of cartilage-specific functional matrix chondromodulin-I mRNA in rabbit growth plate chondrocytes and its responsiveness to growth stimuli *in vitro*. *Biochem. Biophys. Res. Commun.* 249 (3), 885-890.
- Silver, J. (1984). Studies on the factors that govern directionality of axonal growth in the embryonic optic nerve and at the chiasm of mice. *J. Comp. Neurol.* 223:238-251.

- Sims, T.J., and Vaughn, J.E. (1979). The generation of neurons involved in an early reflex pathway of embryonic mouse spinal cord. *J. Comp. Neurol.* 183:707-720.
- Soker, S., Takashima, S., Miao, H.Q., Neufeld, G., and Klagsbrun, M. (1998). Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor. *Cell* 92:735-745.
- Solis, D., Romero, A., Jimenez, M., Diaz-Maurino, T., and Calvete, J.J. (1998). Binding of mannose-6-phosphate and heparin by boar seminal plasma PSP-II, a member of the spermadhesin protein family. *FEBS Letter* 431:273-278.
- Songyang, Z., Fanning, A.S., Fu, C., Xu, J., Marfatia, S.M., Chishti, A.H., Crompton, A., Chan, A.C., Anderson, J.M., and Cantley L.C. (1997). Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science.* 275(5296):73-7.
- Strom, T.M., Nyakatura, G., Apfelstedt-Sylla, E., Hellebrand, H., Lorenz, B., Weber, B.H.F., Wutz, K., Gutwillinger, N., Ruether, K., Drescher, B., Sauer, C., Zrenner, E., Meitinger, T., Rosenthal, A. and Meindl, A. (1998). An L-type calcium-channel gene mutated in incomplete X-linked congenital stationary night blindness. *Nat. Genet.* 19 (3), 260-263.
- Subbaraya, I., Ruiz, C.C., Helekar, B.S., Zhao, X., Gorczyca, W.A., Pettenati, M.J., Rao, P.N., Palczewski, K. and Baehr, W. (1994). Molecular characterization of human and mouse photoreceptor guanylate cyclase-activating protein (GCAP) and chromosomal localization of the human gene. *J. Biol. Chem.* 269 (49), 31080-31089.
- Suzuki, N., Labosky, P.A., Furuta, Y., Hargett, L., Dunn, R., Fogo, A.B., Takahara, K., Peters, D.M.P., Greenspan, D.S., and Hogan, B.L.M. (1996). Failure of ventral body wall closure in mouse embryos lacking a procollagen C-proteinase encoded by *Bmp1*, a mammalian gene related to *Drosophila* tolloid. *Development* 122:3587-3595.
- Takagi, S., Kasuya, Y., Shimizu, M., Matsuura, T., Tsuboi, M., Kawakami, A., and Fujisawa, H. (1995). Expression of a cell adhesion molecule, neuropilin, in the developing chick nervous system. *Dev. Biol.* 170(1):207-22.
- Takahara, K., Lyons, G.E., and Greenspan, D.S. (1994). Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are encoded by alternatively spliced transcripts which are differentially expressed in some tissues. *J. Biol. Chem.* 269:23572-23578.
- Takahara, K., Brevard, R., Hoffman, G.G., Suzuki, N., and Greenspan, D.S. (1996). Characterization of a novel gene product (mammalian tolloid-like) with high sequence similarity to mammalian tolloid/bone morphogenetic protein-1. *Genomics* 34:157-165.

Takahashi, T., Fournier, A., Nakamura, F., Wang, L.H., Murakami, Y., Kalb, R.G., Fujisawa, H., and Strittmatter, S.M. (1999). Plexin-Neuropilin-1 Complexes Form Functional Semaphorin-3A Receptors. *Cell*. 99, 59-69.

Takahashi, T., Nakamura, F., Jin, Z., Kalb, R.G., and Strittmatter, S.M. (1998). Semaphorins A and E act as antagonists of neuropilin-1 and agonists of neuropilin-2 receptors. *Nat. Neurosci.* 1:487-493.

Takito, J., Yan, L., Ma, J., Hikita, C., Vijayakumar, S., Warburton, D., and Al-Awqati, Q. (1999). Hensin, the polarity reversal protein, is encoded by DMBT1, a gene frequently deleted in malignant gliomas. *Am. J. Physiol.* 277(2 Pt 2):F277-89.

Tamagnone, L., Artigiani, S., Chen, H., He, Z., Ming, G., Song, H., Chedotal, A., Winberg, M.L., Goodman, C.S., Poo, M., Tessier-Lavigne, M., and Comoglio, P.M. (1999). Plexins Are a Large Family of Receptors for Transmembrane, Secreted, and GPI-Anchored Semaphorins in Vertebrates. *Cell*. 99, 71-80.

Tartelin, E.E., Kirschner, L.S., Bellingham, J., Baffi, J., Taymans, S.E., Gregory-Evans, K., Csaky, K., Stratakis, C.A. and Gregory-Evans, C.Y. (1999). Cloning and characterization of a novel orphan G-protein-coupled receptor localized to human chromosome 2p16. *Biochem. Biophys. Res. Commun.* 260 (1), 174-180.

Tateno, M., Fukunishi, Y., Komatsu, S., Okazaki, Y., Kawai, J., Shibata, K., Ozawa, Y., Ito, M., Konno, H., Muramatsu, M. and Hayashizaki, Y. (2000). A novel member of SNAG repressor family, mlt 1, is methylated and repressed in mouse liver. (Unpublished).

Thielens, N.M., Bersch, B., Hernandez, J.F., Arlaud, G.J. (1999). Structure and functions of the interaction domains of C1r and C1s: keystones of the architecture of the C1 complex. *Immunopharmacology* 42(1-3):3-13.

Tosi, M., Duponchel, C., Meo, T., and Julier, C. (1987). Complete cDNA sequence of human complement C1s and close physical linkage of the homologous genes C1s and C1r. *Biochemistry* 26:8516-8524.

van Asseldonk, M., Schepens, M., de Bruijn, D., Janssen, B., Merckx, G. and van Kessel, A.G. (2000). Construction of a 350-kb Sequence-Ready 11q13 Cosmid Contig Encompassing the Markers D11S4933 and D11S546: Mapping of 11 Genes and 3 Tumor-Associated Translocation Breakpoints. *Genomics* 66 (1), 35-42.

Velculescu, V.E., Madden, S.L., Zhang, L., Lash, A.E., Yu, J., Rago, C., Lal, A., Wang, C.J., Beaudry, G.A., Ciriello, K.M., Cook, B.P., Dufault, M.R., Ferguson, A.T., Gao, Y., He, T.C., Hermeking, H., Hiraldo, S.K., Hwang, P.M., Lopez, M.A., Luderer, H.F., Mathews, B., Petroziello, J.M., Polyak, K., Zawel, L., Zhang, W., Zhang, X., Zhou, W.,

Haluska, F.G., Jen, J., Sukumar, S., Landes, G.M., Riggins, G.J., Vogelstein, B., and Kinzler, K.W. (1999). Analysis of human transcriptomes. *Nat. Genet.* 23:387-388.

Wechsler-Reya, R.J., and Barres, B.A. (1997). Retinal development: communication helps you see the light. *Current Biology* 7:R433-R436.

Werblin, F.S., and Dowling, J.E. (1969). Organization of the retina of the mudpuppy, *Necturus maculosus*. II Intracellular recording. *J. Neurophysiol.* 32:339-355.

Wheeler, D.L., Chappey, C., Lash, A.E., Leipe, D.D., Madden, T.L., Schuler, G.D., Tatusova, T.A., Rapp, B.A. (2000). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 28(1):10-4.

Wilson, P.J., Meaney, C.A., Hopwood, J.J. and Morris, C.P. (1993). Sequence of the human iduronate 2-sulfatase (IDS) gene. *Genomics* 17 (3), 773-775.

Winberg, M.L., Noordermeer, J.N., Tamagnone, L., Comoglio, P.M., Spriggs, M.K., Tessier-Lavigne, M., and Goodman, C.S. (1998). Plexin A is a neuronal semaphorin receptor that controls axon guidance. *Cell.* 95:903-916.

Wolpert, L., Beddington, R., Brockes, J., Jessell, T., Lawrence, P., and Meyerowitz, E. (1998). *Principles of development*. Oxford University Press, Oxford.

Wozney, J.M., Rosen, V., Celeste, A.J., Mitscock, L.M., Whitters, M.J., Kriz, Hewick, R.M., and Wang, E.A. (1988). Novel regulators of bone formation: Molecular clones and activities. *Science* 242:1528-1534.

Yamada, T., Pfaff, S.L., Edlund, T., and Jessell, T.M. (1993). Control of cell pattern in the neural tube: Motor neuron induction by diffusible factors from notochord and floor plate. *Cell* 73:673-686.

Young, R.W. (1985). Cell differentiation in the retina of the mouse. *Anat. Rec.* 212:199-205.

Young, R.W. (1976). Visual cells and the concept of renewal. *Investigative Ophthalmology and Visual Science* 15:700-725.

Yuan, X., Zheng, X., Lu, D., Rubin, D.C., Pung, C.Y.M., and Sadler, E.J. (1998). Structure of murine enterokinase (enteropeptidase) and expression in small intestine during development. *Am J Physiol.* 274(2) G342-G349.

Zinovieva, R.D., Duncan, M.K., Johnson, T.R., Torres, R., Polymeropoulos, M.H. and Tomarev, S.I. (1996). Structure and chromosomal localization of the human homeobox gene Prox 1. *Genomics* 35 (3), 517-522.

Zou, Y., Stoeckli, E., Chen, H., and Tessier-Lavigne, M. (2000). Squeezing axons out of the gray matter: a role for slit and semaphorin proteins from midline and ventral spinal cord. *Cell*. 102:363-375.

Zuker, C.S. (1994) On the evolution of eyes: Would you like it simple or compound? *Science*. 265:742-743.