BIOCHEMISTRY OF PROX1

by

Tapan P. Patel

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Honors Degree with Distinction in Biochemistry.

Spring 2007

BIOCHEMISTRY OF PROX1

by

Tapan Patel

Approved: _____

Melinda K. Duncan, PhD
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Li Liao, PhD
Committee member from the Department of Computer Science

Approved: _____

Jung-Youn Lee, PhD
Committee member from the Board of Senior Thesis Readers

Approved: _____

John A. Courtright, PhD
Director, University Honors Program

ACKNOWLEDGMENTS

I think I wrote the first draft, but the document you hold now was shaped by many people. Like the cells in your body, almost every sentence of this thesis has turned over at least once, thanks to my mentor and adviser Dr. Melinda Duncan. I have been extremely fortunate to have had the support of Dr. Duncan who took me in her lab as a freshman and over the past four years has constantly pushed me to realize my full potential. I guess the old saying is true that, we strive for greatness mainly because that is what is expected of us from our mentors and colleagues; without people like Dr. Duncan and Dr. Li Liao, I would not have gotten this far.

I actually owe Dr. Duncan a lot more than just her role as my thesis adviser. When I first met her as a freshman, my heart was set on becoming a physician. But over the years, through a series of subtle hints and not-so-subtle conversations, she has slightly and slowly turned me away from medicine and toward a career in scientific research. I guess I finally realized that I was meant to be a researcher when I found myself working in the lab carrying out a Miniprep experiment and programming support vector machines on a Tuesday night at 1AM which also happened to be my 21st birthday. So, Dr. Duncan and Dr. Liao, thank you for that early birthday present; I will cherish it for the rest of my life.

Also I want to thank Xiaoren Chen who taught me all the experimental procedures and has served as my second mentor in the lab; Dr. Jung-Youn Lee for providing helpful comments on this thesis; and Dr. Liao who showed me the joy of

bioinformatics and fostered my curiosity in computational biology for the past year. I also want to thank Dr. Sandeep Patel and Dr. Brian Bahnson for their advice on the structural modeling aspect of my thesis. Also I am grateful to the entire Department of Chemistry and Biochemistry for granting me access to the Chemistry Cluster computers where all the structure predictions and analyses were done.

Furthermore, none of my current work would be possible without the generous financial support from the Arnold and Mabel Beckman Foundation, the Barry M. Goldwater foundation, University of Delaware Undergraduate Research Program, the Howard Hughes Medical Institute, and NEI EY012221 grant (MKD).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Vertebrate development is a highly orchestrated process, coordinated by thousands of proteins, and involving hundreds of pathways that specifically and selectively allow stage-wise transition from a fertilized embryo to an adult individual. In this grand maze, Prox1 stands out as one of the few key transcription factors that influences control over various stages of development, including the development of the liver, pancreas, lens, and the lymphatic system. It has also been implicated to play a role in cancer progression; however the current research is only speculative. Although the expression pattern of this multifunctional homeodomain transcriptional factor Prox1 has been established, the function of Prox1 and more importantly the functional regulation of Prox1 in different cell types are both poorly understood in molecular detail. In general, functional regulation of proteins can be grouped in three broad categories: transcriptional control, posttranslational modification or protein-protein interactions. Here we report how the function of Prox1 is critically modulated by interaction with specific protein partners. In particular, we have found that the splicing factor sc35 directly interacts with Prox1 and represses its function as a transcription factor. However, since sc35 is normally localized to the interchromatin granule (nuclear speckles) regions, Prox1-sc35 interaction is only possible during times of high transcriptional activity. Thus, we propose that sc35 functions to keep Prox1 mediated gene activation under control. In addition to sc35, other proteins such as proliferating cell nuclear antigen, PCNA, also specifically interacts with Prox1.

Structure prediction of the Prox1-PCNA protein complex has allowed us to propose a mechanism whereby monomeric PCNA molecule strongly interacts with Prox1 in non-proliferating cells and represses its function as a transcription factor. Since Prox1 is mainly involved in cell differentiation which typically prerequisites/accompanies cellular proliferation, the PCNA-Prox1 interaction might coordinate the timing of differentiation and proliferation during development. The relative amounts of Prox1 and PCNA and the specific Prox1-PCNA interaction could also play a role in the decision to differentiate or proliferate. These findings strongly suggest that protein-protein interactions play a significant role in the functional regulation of Prox1 during normal vertebrate development.

Chapter 1

INTRODUCTION

Development of mammalian systems is well understood at the morphological level but is poorly understood overall at the molecular level. Although we know of a vast number of proteins and transcription factors that are involved in the regulation and cell fate determination that ultimately shape an organism, the way these pieces of the puzzle fit together is a work in progress. And, perhaps it will be decades before the entire picture is clear, but in order to form the complete picture that explains how one goes from an embryo to an adult being, we undoubtedly need to understand the role of as many pieces as we can.

In the vast pool of proteins, a small subset of transcription factors are particularly important for the development of many organ systems. These "master regulators" are especially good candidates for further study. Prox1 is one such transcription factor that is involved in many different developmental processes including development of the liver, pancreas, lens, and retina as well as the lymphatic system (Wigle JT, Chowdhury K et al. 1999; Wigle JT and G 1999; Burke 2002; Duncan MK, Cui W et al. 2002; Hong YK, Harvey N et al. 2002; Dudas J, Papoutsi M et al. 2004; Wang J, Kilic G et al. 2005) . Although Prox1 has been identified in these organs, its role and more importantly, the biochemistry of Prox1, that is, how its

function is modulated, has not yet been addressed. This work aims to elucidate how Prox1 function is modulated in different cell types.

1.1 Prospero homolog of Prox1

The vertebrate Prox1 protein has a *Drosophila* homolog, Prospero, which has been studied extensively. Prospero, like Prox1, is a homeobox transcription factor and as such, it serves many different functions in *Drosophila*. Primarily, Prospero is responsible for neuroblast cell fate determination and for regulating the gene expression of rhodopsin. During neurogenesis, neuroblast stem cells (denoted as NB in fig 1) divide asymmetrically to give rise to two daughter cells – a ganglion mother cell and a neuroblast stem cell (Doe 1991; Doe 1992; Jan 1998; Myster 2000). The daughter neuroblast stem cell repeats this asymmetric division whereas the ganglion mother cell divides once more to generate two daughter cells which differentiate into neurons or glia (Jan 1998). The asymmetric division is due to asymmetric distribution of cytoplasmic Prospero in neuroblast stem cells. *Prospero* RNA and protein have been found to be localized to the cytoplasmic membrane of neuroblasts which enables the *Prospero* to segregate into only one daughter cell, namely, the ganglion mother cell (Hirata 1995; Knoblich 1995; Spana 1995). Once in the ganglion mother cell, Prospero translocates from the cytoplasm into nuclei and initiates the differentiation of the ganglion mother cell into neurons or glia by activating or repressing certain cell

type specific genes. In this regard, Prospero is thought to force the ganglion mother cell to exit the cell cycle and initiate the neural differentiation program.

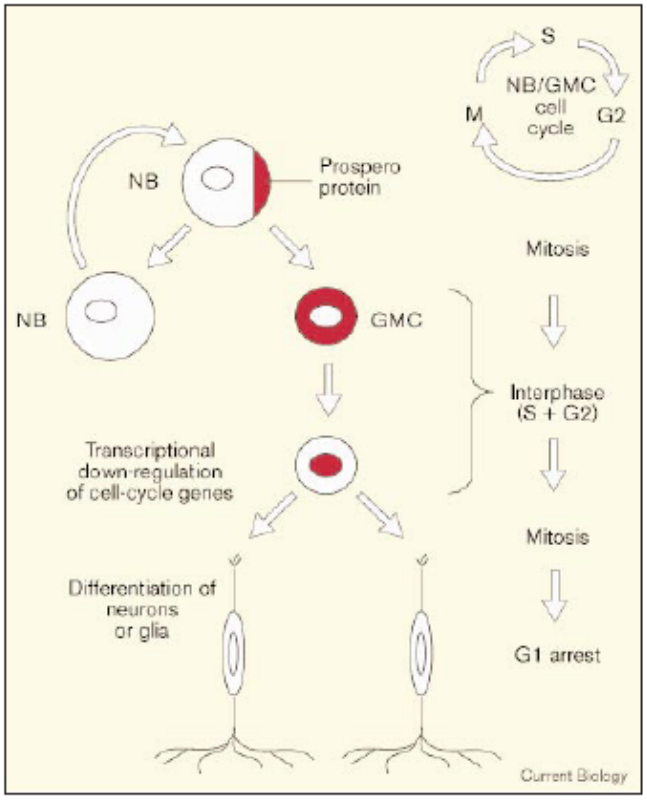


Figure 1: The role of Prospero in neuroblast cell fate determination. Prospero protein (in red) is synthesized in neuroblast stem cells and is localized to the cell cortex. Upon division of NB, Prospero is asymmetrically distributed solely to the ganglion mother cell (GMC) where it migrates to the nucleus and enacts a transcription program leading to neural differentiation (Duronio 2000).

Prospero is also found to be regulated by the cell cycle regulatory protein Cyclin E. As discussed above, asymmetric distribution of Prospero in neuroblast cells results in glial cell fate. However, loss of Cyclin E causes an even distribution of Prospero and induces all neuroblast cells to differentiate into glial cells (Berger 2005).

This direct link between Cyclin E and Prospero helps explain the role of Prospero in cell fate specification.

A similar situation occurs with Prox1 in vertebrate development as discussed below.

1.2 Prox1 Structure

Prox1, like Prospero, is a homeobox transcription factor. Human *prox1* gene encodes a protein of 737 amino acids with an estimated molecular weight of 84 kDa. The protein consists of C-terminal divergent homeodomain and a novel Prospero domain, an N-terminal nuclear localization signal and three nuclear receptor boxes (Burglin 1994; Tomarev 1998). Prox1 and Prospero are most similar in the C-terminal domains and in fact, the homeo-Prospero domain (HDPD) of Prox1 is conserved among most vertebrates. The domain structure of Prox1 is illustrated in Fig 2 and a multiple sequence alignment of the C-terminal residues of Prospero/Prox1 protein in fig 3.

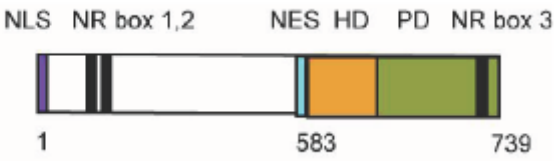


Figure 2: Prox1 contains an N-terminal nuclear localization signal (NLS), a C-terminal Homeo domain (HD) and Prospero domain (PD) and a nuclear export signal (NES), along with three nuclear receptor boxes (NR 1-3).

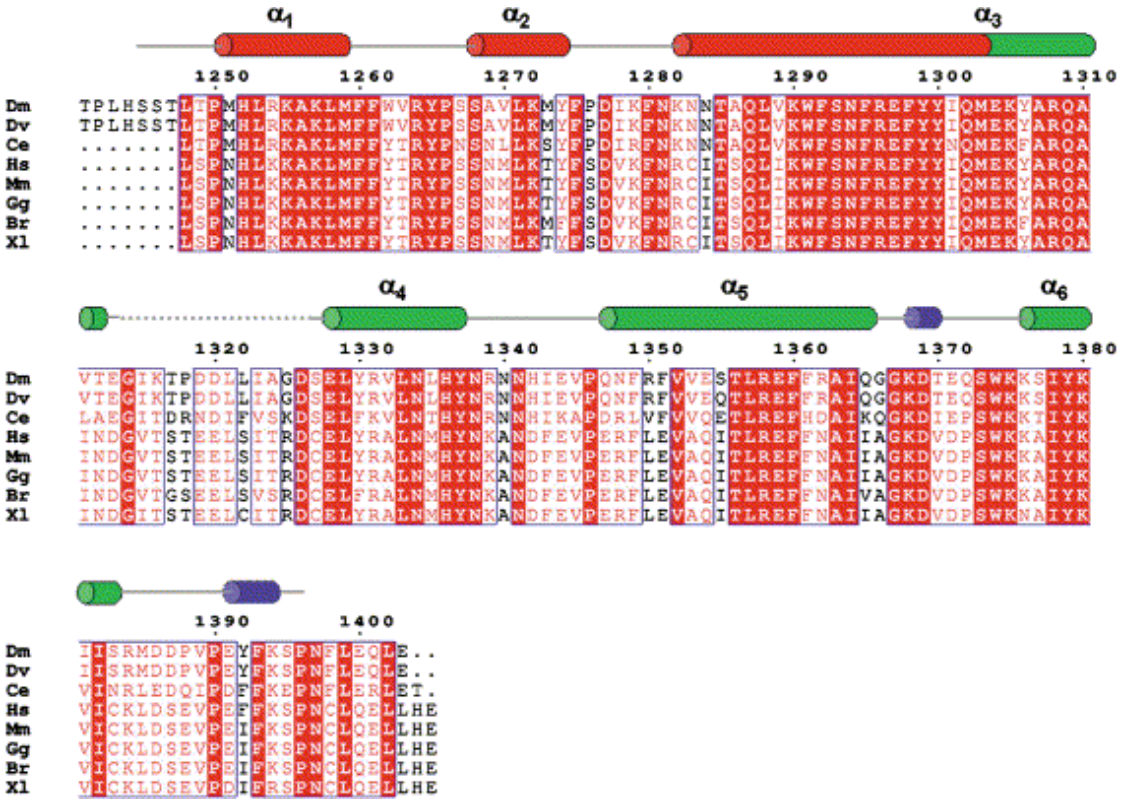


Figure 3: Multiple sequence alignment of the C-terminal residues of Prox1/Prospero protein from various species show significant conservation (highlighted in red). The alpha helices are labeled α_1- α_6 with red labels corresponding to the homeodomain and green Prospero and 3_{10} helices in blue. The sequences include about 160 amino acids from *Drosophila melanogaster* (Dm, residues 1241–1403), *Drosophila virilis* (Dv, residues 1394–1556), *Caenorhabditis elegans* (Ce, residues 430–586), *Xenopus laevis* (Xl, residues 583–740), zebrafish (Br, residues 582–739), chicken (Gg, residues 579–736), mouse (Mm, residues 580–737), and human (Hs, residues 579–736; numbering for *D. melanogaster* shown). (Jodi M. Ryter 2002).

The DNA binding ability of Prox1 is entirely carried out through the HDPD. The homeodomain is a common structural motif found in many transcription factors. It is a highly conserved DNA-binding domain made up of approximately 60 amino acid residues, and in fact, most of what we know about protein-DNA interactions comes from studies on homeodomain proteins. Some notable proteins in the homeodomain family include the Hox gene cluster, and *S. cerevisae* mat a/alpha

gene. A homeodomain is characterized as containing three alpha helices, two of which lie parallel to each other and the third lies across from the two thus creating a helix-loop-helix-turn-helix fold. The third helix contains a DNA recognition site and lies in the major groove of DNA where it makes sequence specific contacts with a double stranded DNA molecule. The loop connecting the helices I and II also makes contacts with the phosphate backbone of DNA and further stabilizes the interaction (Gehring 1994; Yousef 2005). Although the homeodomain of Prox1 does not share significant sequence similarity with other homeodomain family members, threading analysis of Prox1 C-terminal sequence has shown that the Prox1 homeodomain is still capable of adopting the three helix-bundle motif (Banerjee-Basu 1999). This prediction has later been confirmed by the three-dimensional structure of Prospero in which Prospero adopts an atypical homeodomain configuration (the structure of Prox1 has not yet been solved) (Jodi M. Ryter 2002; Yousef 2005). Furthermore, as already mentioned, Prox1 also contains a novel Prospero domain with unknown function. The Prospero domain region is composed of a four-helix bundle with anti-parallel consecutive helices. The three-dimensional structure of the C-terminus of the Prospero protein reveals a plausible role for the Prospero domain. Almost all known homeodomains contain a DNA recognition helix at the extreme C-terminus of the protein; however, the striking difference in Prox1 and Prospero proteins is that the DNA recognition helix (α_3) connects the putative homeodomain and Prospero domains together as one structural motif. The 3-D structure of Prospero reveals that

the Prospero domain can make further contacts with DNA and also functions to mask the nuclear export signal (NES). It is believed that upon DNA binding, a slight conformational change positions the Prospero domain in front of NES (Yousef 2005). The overall fold of Prospero and its interaction with a non-canonical DNA helix is depicted in figure 4.

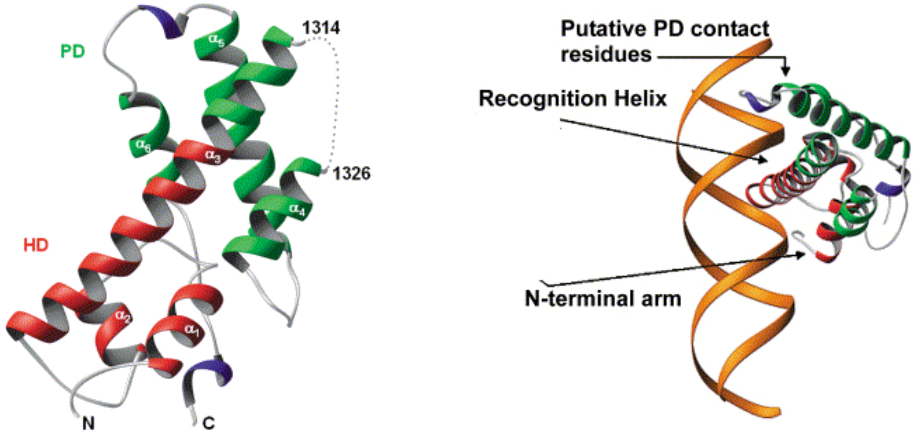


Figure 4: (left) Overall fold of Prospero protein. Homeodomain (HD) in red and Prospero domain in green are connected as one structural by the DNA recognition helix α_3. (right). DNA bound structure of Prospero. Helix III is inserted in the major groove of DNA where it makes sequence specific contacts (Jodi M. Ryter 2002).

1.2 Role of Prox1 in development

Prox1 is expressed in ocular tissues (retina, lens, and cornea), the lymphatic system, brain, liver, pancreas, ovary and testis, cochlea, skeletal muscle and heart. Its role in several of these organs is now discussed.

1.2.1 Prox1 in lens development

Mouse lens development has been studied extensively, and it is very similar to human lens development. The whole process of lens development can be grouped into three basic steps: bias, specification and differentiation. In the bias phase, the endoderm and mesoderm interact with the adjacent prospective head ectoderm to give the head ectoderm a lens-forming bias. The optic vesicle then extends from the diencephalon to set up the proper spatial relationship between the lens and the retina. When the optic vesicle meets the head ectoderm, it induces the formation of a lens placode, which then invaginates to form the lens vesicle (Piatigorsky 1981; Saha MS 1989; Grainger 1992). At this stage, the lens vesicle is just a hollow ball of cells.

The differentiation of the lens tissue into a transparent structure that is capable of directing light onto the retina involves a series of changes in both cell structure and shape as well as the synthesis of lens-specific proteins called crystallins (Fig 5). The posterior cells of the lens vesicle elongate and, under the influence of the neural retina, become the lens fibers. These elongating cells form the primary lens fibers and as they continue to grow, they synthesize crystallins, which eventually fill up the cell. At the same time, the organelles of the lens fiber cells undergo programmed degradation. The anterior cells of the lens vesicle constitute a germinal epithelium, which keeps dividing. These dividing cells move toward the equator of the vesicle, and as they pass through the equatorial region, they, too, begin to elongate, forming concentric circles of secondary lens fiber cells (Piatigorsky 1981).

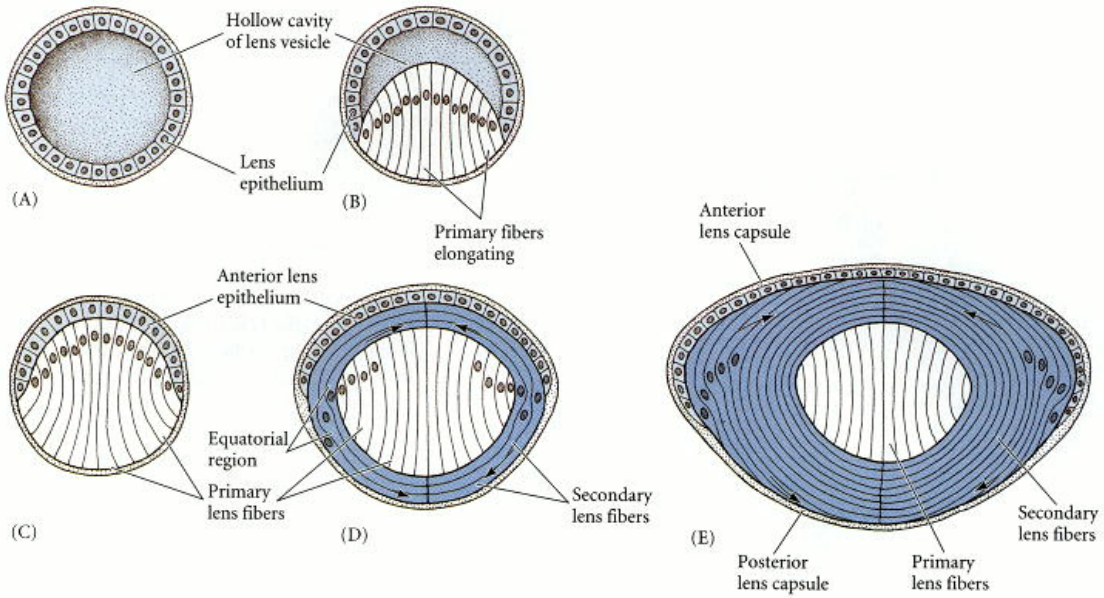


Figure 5: Differentiation of the lens cells. (A) Lens vesicle as a hollow ball of cells. (B, C) Fiber cells elongate anteriorly to fill the hollow lens vesicle. (D) New secondary fiber cells derived from the anterior lens epithelium. (E) Lens consisting of postmitotic differentiated fiber cells expressing crystallins, and a monolayer of proliferating epithelial cells in the anterior portion of the lens (Craig 1974).

Obviously this highly controlled process is regulated by various growth factors and transcription factors. The decision to differentiate posterior cells and proliferate anterior cells is thought to be the result of combination of FGF (fibroblast growth factor) signaling and the activity of certain transcription factors including PAX6, Mafs, Sox, and Prox1. In the lens vesicle, there is an FGF gradient such that low concentration of FGF in the anterior region promotes epithelial cell proliferation and a higher concentration in the posterior side induces epithelial cell elongation and fiber cell differentiation (McAvoy 1989). Along with FGF signaling, Pax6 plays key roles in lens bias and lens specification. Pax6 is a multifunctional transcription factor with a paired homeodomain. Its importance in lens specification became clear when

elimination of Pax6 protein by conditional knockout from the ectoderm after the lens bias stage and during the lens specification stage resulted in the complete absence of all lens structures (R. Ashery-Padan 2000). Furthermore, the lens-specification marker Sox2 was not expressed in $Pax6^{-}/Pax6^{-}$ embryos (Hogan 1998). Lens-bias is achieved by the activation of Sox2 in the ectoderm by Pax6. During lens differentiation, Pax6 controls the expression of the homeobox genes *Six3* and *Prox1*. During this stage, Pax6 and Sox2 bind cooperatively to the δcrystallin enhancer and drive its gene expression in the lens placode. The transcriptional activation is upregulated by the basic leucine zipper Maf transcription factor. The interaction of these key transcription factors is depicted in figure 6.

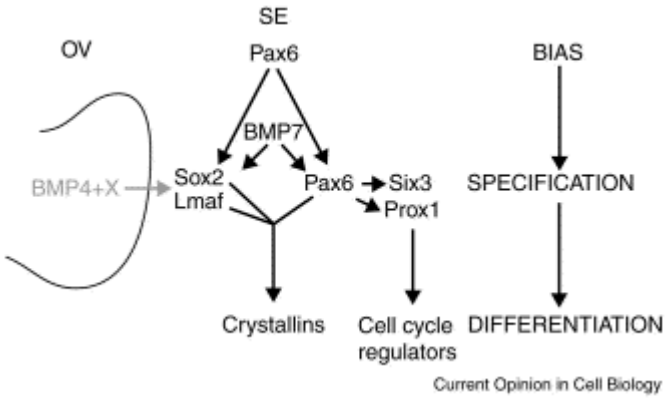


Figure 6: Transcriptional regulation of lens development in vertebrates. Early expression of Pax6 is required for the upregulation of Sox2 and for maintaining high levels of Pax6. Growth factors secreted from the ocular vesicle (OV) further increase Sox2 and the basic leucine zipper Lmaf protein expression. During specification stage, Pax6 drives the expression of Prox1 and Six3 and a combination of Pax6, Sox2, and Lmaf is required to trigger the expression of lens structural proteins such as crystallins (Gruss 2001).

The transcriptional regulation outlined here draws parallel with *Drosophila* eye development to which the reader might be more familiar. Pax6 plays a similar role in vertebrate eye development as the ey (eyeless) protein in *Drosophila*. The Ey protein binds directly to numerous genes involved in eye development including the lens crystallin genes. In fact, in a famous experiment (by Walter Gehring) where the *ey* gene was artificially expressed in cells that normally give rise to leg tissue, a misplaced eye developed on the leg (Halder G 1995). Similarly, *Sox2* has the counterpart *so* (*sine oculis*) and *eya* (*eyes absent*). These proteins work in positive feedback loop, much like in vertebrate lens development.

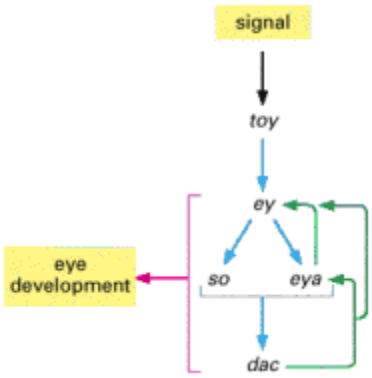


Figure 7: Transcriptional regulation of eye development in *Drosophila*. Ectopic expression of either *toy* (twin of eyeless) or *ey* (eyeless) triggers the transcription of *so* (*sine oculis*) and *eya (eyes absent)* genes leading to eye formation. Green arrows indicate positive feedback loop (Czerny T, Halder G et al. 1999).

Although Pax6 is a key transcription factor for the early stages of lens development, the later differentiation stage is heavily dependent on Prox1 activity which is mainly involved in cell cycle regulation. In the developing eye, Prox1 is first

expressed in the lens placode, the lens vesicle, and in the anterior epithelium and fiber cells (Wigle JT, Chowdhury K et al. 1999; Duncan MK, Cui W et al. 2002). Prox1 inactivation in knockout mice causes abnormal cellular proliferation, downregulated expression of the cell-cycle inhibitors p27KIP1 and p57KIP2, misexpression of E-cadherin and inappropriate apoptosis. These abnormalities result in unpolarized and non-elongated epithelial cells leading to a hollow lens. Thus, fiber cell differentiation and elongation is dependent on Prox1 activity.

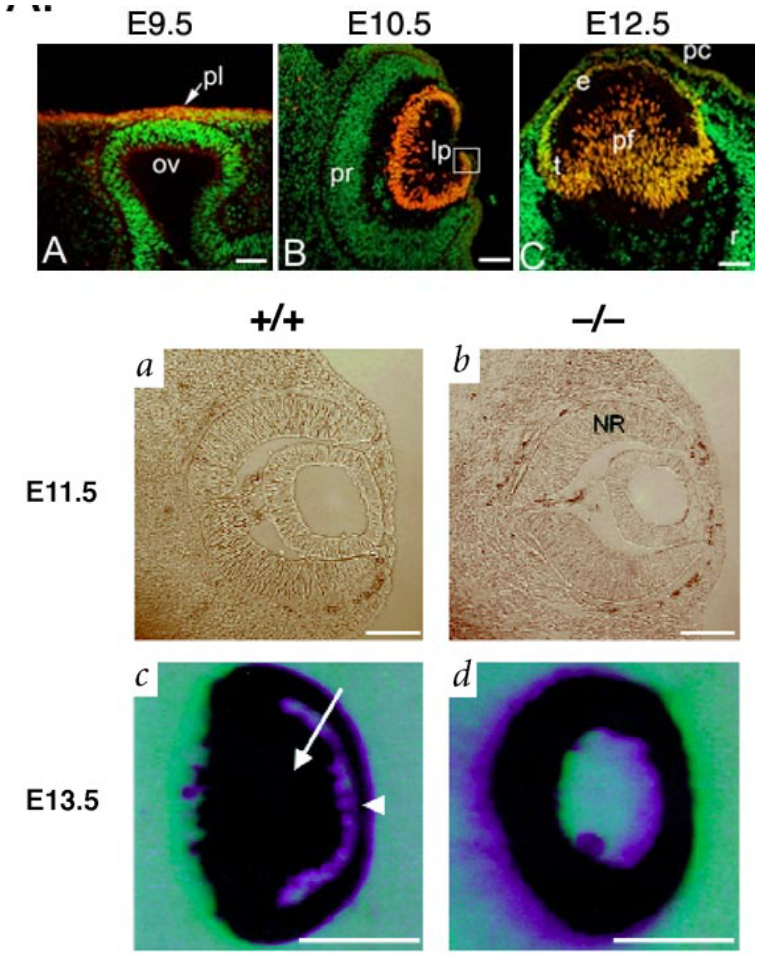


Figure 8: (top) Prox1 expression pattern in developing mouse lens. At E9.5, Prox1 is expressed in the lens placode and by E10.5, it is found in the lens vesicle. By E12.5, Prox1 is highly expressed in lens fiber cells (Duncan MK, Cui W et al. 2002). (bottom) Lens fiber elongation is compromised in Prox1$^{-/-}$ mice. Normal lens development in wild-type E11.5 and E13.5 embryos is shown as Nomarski imaging (a) and LacZ staining (c). In Prox1$^{-/-}$ mice, the lens vesicle is formed normally; however, the lens-fibre elongation is absent as evidenced by the hollow lens (Wigle JT, Chowdhury K et al. 1999).

1.2.2: Role of Prox1 as a transcription factor in lens development

Aside from its role in lens fiber cell elongation, Prox1 acts as a true transcription factor by turning on the expression of βB1-crystallin genes in developing chicken lens. Crystallins are water-soluble proteins that make up the majority of

protein content of lens. They are grouped into three major types: alpha, beta, and gamma crystallins. As already mentioned, Pax6, Sox2 and Lmaf are all involved in the transcriptional regulation of the gamma-crystallin genes. Similarly, Prox1 binds specifically to the OL2 element of the βB1-crystallin promoter region and activates downstream gene expression. A remarkable feature of the lens is that almost all crystallin protein is found in the lens fiber cell and none is synthesized by the epithelial cells. A possible explanation for this comes from work in our lab. Mafs, Prox1, and Pax6 can all bind to three important *cis* elements, PL1, PL2 and OL2 in the βB1-crystallin promoter. Pax6 however, is a repressor of the βB1-crystallin promoter and binds to all three elements. Transfection assays showed that in the absence of Pax6, Mafs and Prox1 activate the βB1-crystallin promoter; however, addition of Pax6 (in the presence of Mafs and Prox1) represses the βB1-crystallin promoter. Thus, our lab proposed the model that Pax6 occupies the three PL1, PL2 and OL2 elements and represses the βB1-crystallin promoter in lens epithelial cells. Displacement of Pax6 by Mafs and Prox1 in lens fiber cells then activate the promoter (Wenwu Cui 2004). This modulation of Prox1 function by specific protein-protein interactions will become the central theme of this thesis.

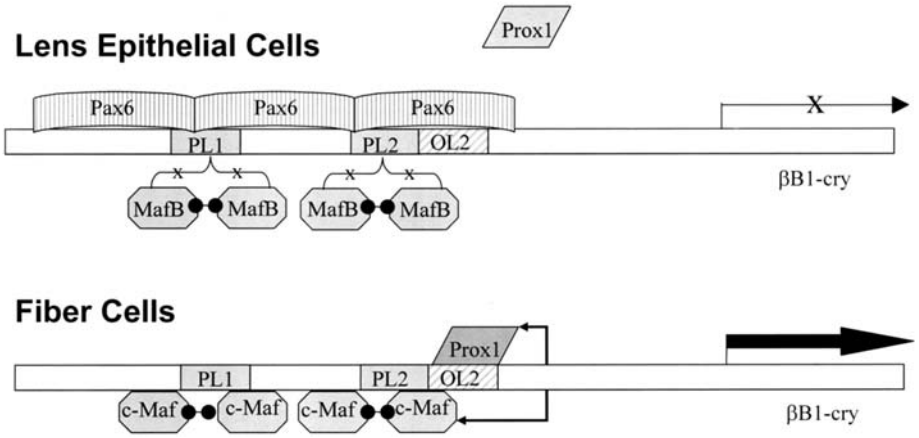


Figure 9: Model for the transcriptional regulation of βB1-crystallin gene. High levels of Pax6 in the lens epithelial cells repress the βB1-cry promoter. In lens fiber cells, lower levels of Pax6 allow displacement by Mafs and Prox1 which activate the gene expression (Wenwu Cui 2004) .

Prox1 also activates the γF-crystallin gene (Lengler J 2001) and has been shown to activate the chicken αA, mouse αA, and chicken δ1 crystallin genes in PLEs (MK Duncan unpublished data).

Taken together, Prox1 plays many key roles in normal lens development.

1.2.3: Role of Prox1 in lymphangiogenesis

Our body has two distinct vascular networks, one composed of blood vessels and the other composed of lymphatic vessels. The blood vessels deliver important nutrients and carry immune cells throughout the body. Whereas, the primary function of the lymphatic system is to return extravasated immune cells as well as fluid leaked from the blood vessels back into the blood vascular system for circulation. In bird and mammals, the lymphatic system is also important in immune response and absorption of fat from the gut. These two plumbing systems of our

bodies originate from a common parent and separate during embryonic development. Abnormalities in the separation of these two systems lead to severe consequences such as human congenital diseases involving arteriovenous malformations, edema, and accumulation of fluid in the peritoneal cavity (Turner M 1995; Farhad Abtahian 2003).

Lymphatic vessels primarily originate from blood vessels (Carmeliet 2002; Oliver G and M. 2002). During embryogenesis, endothelial cells of early embryonic veins (cardinal vein) begin to express the receptors LYVE-1 and VEGFR-3. An unidentified signal then triggers polarized expression of Prox1 such that different regions of the cardinal vein express different amounts of Prox1. The polarized expression of Prox1 marks the first committed step towards a lymphatic lineage (step 2 in the figure below). Cells expressing LYVE-1, VEGFR-3 and Prox1 begin to form buds and start to produce lymphatic endothelial cell specific markers such as SLC and increased production of VEGRF-3. These cells have now become dedicated to lymphatic fate. On the other hand, the polarized expression of Prox1 results in some cardinal vein endothelial cells that do not contain Prox1. In these cells, the VEGFR-3 and LYEV-1 production is decreased leading to a blood vessel lineage. In this regard, Prox1 is the master regulator for lymphangiogenesis. Prox1 null mice show little sign of budding and the lymphatic specific markers VEGFR-3, LYVE-1 and SLC are no longer present either; although vasculogenesis and angiogenesis of the vascular system is unaffected. Furthermore, over-expression of Prox1 was sufficient to

induce lymphatic specific gene expression in blood endothelial cells while inhibiting blood specific gene expression (Wigle JT and G 1999; Hong YK, Harvey N et al. 2002; Tatiana V. Petrova 2002).Taken together, these data suggest that Prox1 is a specific and required regulator for the development of the lymphatic system and that the vascular and lymphatic systems develop independently as depicted in the figure below.

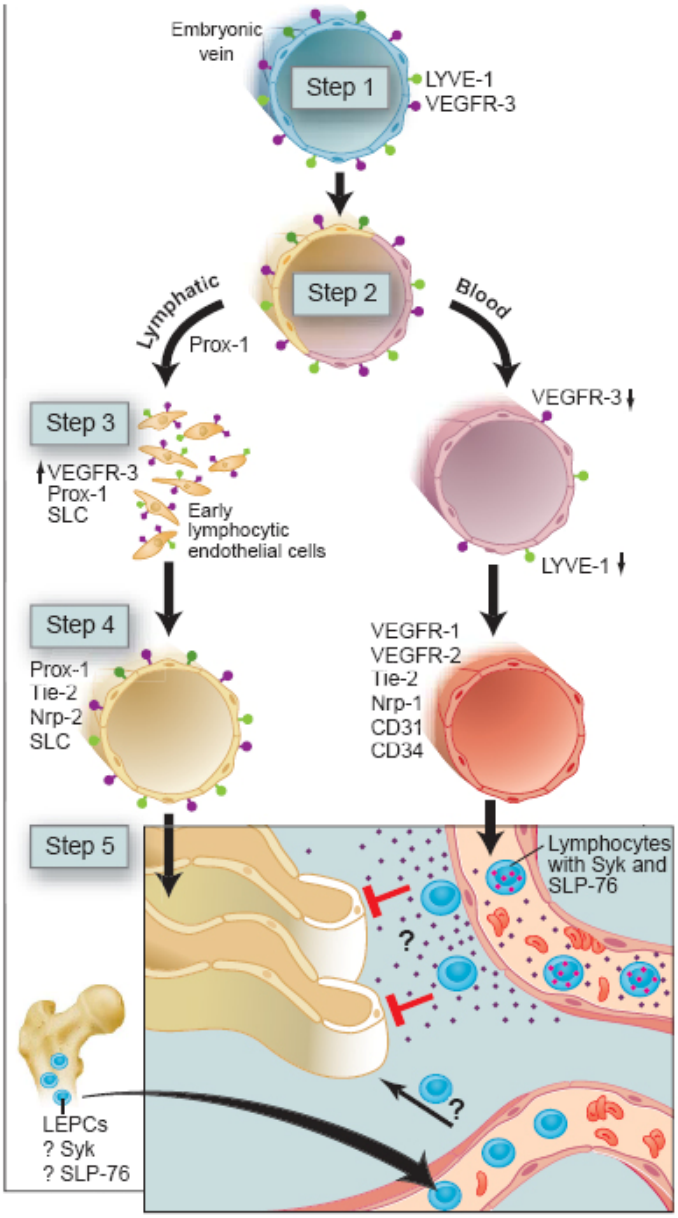


Figure 10: Lymphatic and blood vessels original from common embryonic veins. Polarized expression of Prox1 in a subpopulation of endothelial cells leads to budding and sprouting that gives rise to the lymphatic system. Prox1 absent cells give rise to blood vessel lineage (Padera 2003).

Notice the similarity between lymphangiogenesis and neuroblast cell fate determination in *Drosophila* where polarized expression of Prospero leads to the differentiation of neural stem cells.

1.2.4: Role of Prox1 in liver and pancreas development

The vertebrate liver is derived from both, the endoderm of the ventral foregut and the mesenchyme of the septum transversum. Under the induction of the cardiac mesenchyme, the foregut endoderm forms the liver bud from which early hepatocytes migrate in a cord-like fashion into the mesenchyme of the septum transversum. In vertebrates, the primary hepatic rudiment is an endodermal evagination of the ventral foregut that extends into the mesenchyme of the septum transversum. After hepatic-bud formation, the basal membrane surrounding the hepatic epithelium degrades. Cords of hepatocytes emerge from the hepatic epithelium, invading the septum transversum and lateral mesenchymal areas of the hepatic lobes (Zoe¨ Burke 2002).

Both liver and pancreas arise from an early endoderm with a bipotential liver and pancreas precursor cell type. The expression of Prox1 in restricted regions of the early endoderm then gives rise to the mammalian liver and pancreas. In fact, Prox1 is one of the earliest markers of the liver cell type. It starts to express in the hepatic endoderm at 8.5dpc in mouse. Later it is detected in the hepatic bud, gall bladder and dorsal and ventral pancreatic primordial (Sosa-Pineda 2000; Dudas J, Papoutsi M et

al. 2004). Although Prox1 is not necessary for initiating liver development, it is required for the migration of the hepatocytes into the septum transversum (Sosa-Pineda 2000). Homozygous Prox1 knockout mice show noticeable reduction in the liver size due to the absence of hepatocytes from most of the liver lobes, although the overall formation of liver was not impaired (Burke 2002).

Prox1 is also highly expressed in pancreas and plays important role in its development. The mouse pancreas begins to form at day 9.0dpc at the foregut/midgut region. The growth of pancreas is characterized by the formation of distinct pancreatic cells types including endocrine (alpha, beta, delta, PP), exocrine and ductal cells. These cells appear in well-defined sequential order: glucagon-producing alpha cells appear first, followed by insulin-producing beta cells, then pancreatic exocrine cells, somatostatin-producing delta cells and pancreatic polypeptide-producing (PP) cells come last (Slack 1995; Murtaugh 2003). Prox1 starts to express at dorsal region at day 9.5dpc and is widely expressed in all pancreas progenitor cells. Later its expression is restricted to endocrine cells, the duct but not pancreatic exocrine cells.

Similar to liver development, Prox1 knockout severely affects the size and morphology of the pancreas, although the initiation of pancreas formation is unaffected. Furthermore, endocrine cell formation was decreased and exocrine cell formation increased, suggesting that Prox1 is essential for cell fate determination in developing pancreas (Wang J, Kilic G et al. 2005).

1.3 Plausible role of Prox1 in cancer progression

Prox1 has been shown to be crucial for the formation of new lymphatic vessels and for cell fate determination in various organs by exercising control over cell cycle regulatory proteins. These processes are extremely important for homeostasis and are closely tied with cancer progression. It is clear that most cancer types involve some loss of cell cycle control whereby cells either replicate erroneously or become apoptotic. Furthermore, a benign cancerous cell cannot metastasize without plenty of nutrients and good blood flow. Thus angiogenesis, formation of new blood vessels, is crucial for cancer progression. These two important aspects of cancer progression are to some extent controlled by Prox1. As discussed above, high levels of Prox1 in endothelial cells induces the formation of lymphatic vessels, lymphangiogenesis, whereas low levels are accompanied by formation of blood vessels, angiogenesis. So it is plausible that angiogenesis (necessary for cancer metastasis) can be inhibited by increased Prox1 expression. Furthermore, loss of cell cycle control and abnormal differentiation upon Prox1 deletion implies that Prox1 may function as a tumor suppressor gene. For example, the lens of Prox1 knockout mice show defects in lens fiber cell differentiation, abnormal cellular proliferation (hallmark of cancerous cell type) together with down-regulation of cell cycle inhibitors p27 (Kip1) and p57 (Kip2) (Wigle JT, Chowdhury K et al. 1999).

The possibility of Prox1 being a tumor suppressor has been validated by the finding that some hematologic malignant cell lines contain a silenced copy of

Prox1 due to hypermethylation of Prox1 DNA at intron 1, similar to what has been observed in primary lymphomas (Nagai H, Li Y et al. 2003). Furthermore, Prox1 is found to be either absent or drastically reduced in 63% of bilary tract carcinomas. Both in hematologic malignancies and bilary tract carcinomas, similar mechanisms like genomic deletions and hypermethylation, which are prototypic for the inactivation of tumor suppressor genes, also inactivate *Prox1* (Laerm A, Helmbold P et al. 2007).

1.4: Regulation of Prox1 by other proteins

It was reported in section 1.2.2 that Prox1 is able to selectively turn on the βB1-crystallin promoter in developing lens by binding to the OL2 element of the promoter region. However, as was pointed out, βB1-crystallin promoter is only turned on in the differentiated lens fiber cells and not in the lens epithelial cells. This is partly because high levels of Pax6 in the epithelial cells inhibit the βB1-crystallin promoter by not allowing Prox1 access to the OL2 element. Decreased levels of Pax6 in the lens fiber cells then allows Prox1 to bind to the OL2 element and turn on βB1-crystallin promoter. So it is already clear that the function of Prox1 can be easily modulated by interaction with other proteins and in fact, the protein-protein interactions might be crucial to Prox1 function as a transcription factor. Further evidence in favor of this hypothesis comes from the interaction between Prox1 and several orphan nuclear receptors. Prox1 works as a transcriptional co-repressor to the human cholesterol 7alpha-hydroxylase (*CYP7A1)* gene in HepG2 cells through direct interactions with

orphan nuclear receptors, hepatocyte nuclear factor 4alpha (HNF4a, NR2A1) and liver receptor homolog 1 (LRH-1, NR5A2). CYP7A1 is an important enzyme for maintaining lipid homeostasis as it catalyzes the first and rate-liming step in the conversion of cholesterol to bile acids (Jun Qin 2004; Steffensen KR, Holter E et al. 2004; Song KH, Li T et al. 2006). Both HNF4a and LRH-1 regulate lipid metabolism by controlling *CYP7A1* gene expression. As was noted in section 1.1, Prox1 contains three nuclear receptor (NR) boxes, two at the N-terminus and one at the C-terminus. The N-terminal NR box I is found to interact with HNF4a and LRH-1 resulting in the repression of *CYP7A1* gene (Yi-Wen Liu 2003; Jun Qin 2004; Steffensen KR, Holter E et al. 2004; Song KH, Li T et al. 2006).

Prox1 also interacts with the LRH-1 homolog steroidogenic factor 1 (SF-1) through NR box I. SF-1 is a multifunctional nuclear receptor protein. It is a key regulator during steroidogenic tissue development and controls the expression of all steroidogenic enzymes as well as the expression of cholesterol transporters and steroidogeneic-stimulating hormones. SF-1 is also a regulator of genes involved in sex determination (Keith L. Parker 2002; Pierre Val 2003). The interaction of Prox1 with SF-1 results in inhibition of SF-1 function (i.e. Prox1 inhibits SF-1 driven gene transcription). This inhibition is actually crucial for inter-renal organ development (Yi-Wen Liu 2003).

Thus, Prox1 is truly a multifunctional transcription factor. It can function as a classical transcription factor by binding to promoter regions of a gene and

activating or repressing gene transcription, or as we have just seen, Prox1 can also interact with other transcription factors (specifically nuclear receptors) and change the function of the interacting partner thereby exercising indirect control over gene transcription. **This functional modulation upon specific protein-protein interactions is particularly interesting as it provides a plausible explanation of how a single transcription factor can affect the development of so many different processes. For example, depending on the protein pool in a given cell, Prox1 may take on different roles in different cell types. Thus, in order to fully understand the biochemistry of Prox1 function, we aim to identify the protein interacting partners of Prox1 and study the consequences of such interactions on Prox1 function.**

Chapter 2

MATERIALS AND METHODS

2.1 Immunohistochemistry

2.1.1 Immunofluorescent labeling of tissue sections for Prox1

Human lenses were obtained from a 21 year old accident victim via the Oregon Lions Eye Bank. Eyes were isolated from 10 day embryonic *Gallus gallus* (chicken), 5 month old *Rattus rattus* (Sprague Dawley rat), adult *Rana pipens* (grass frog), adult *Anolis sagrei* (lizard) and adult *Fundulus similes (*fish). The unfixed tissue was placed in OCT (Tissue Tek) and frozen on a bed of dry ice. Immunohistochemical staining was performed as previously described (Reed NA 2001). Briefly, tissue was sectioned at 16 μm and mounted onto Colorfrost/Plus slides (Fisher Scientific, Pittsburgh, PA). The slides were fixed with ice cold 1:1 acetone:methanol for 10 minutes and blocked with 1% BSA in 1X PBS for 1 hour at room temperature. Hybridoma supernatants were diluted in 1% BSA-PBS, layered over the sections and incubated for one hour at room temperature. The slides were washed twice with 1X PBS and then incubated for one hour at room temperature with a 1:250 dilution of goat-anti mouse Alexa Fluor 568 conjugate (Molecular Probes, Eugene, OR) in 1% BSA/PBS containing a 1:2000 dilution of the nucleic acid stain

To-Pro3 (Molecular Probes). The slides were again washed twice with 1X PBS and then mounted. The slides were viewed on a Zeiss 510 LSM confocal microscope with an Argon/Krypton laser (Zeiss, Gottingen, Germany).

2.1.2 Immunofluorescence of Prox1 and SC35 in cultured cells

Cells were plated in 4-well Lab-TEK II chambers (Nalge Nunc, Napeville, IL) at a density of 50,000 cells per well and grown overnight. The next day, cells were washed with 1X PBS and fixed with 4% paraformaldehyde (Fisher Scientific) for 10 minutes. After rinsing with 1X PBS, cells were permeabilized with TritionX-100 (0.2%, Sigma) for 5 minutes. The cells were rinsed gently with 1X PBS twice and blocked with blocking solution (3% BSA in 1X PBS) for 30 minutes with shaking in room temperature. 5G10 ascites and SC35 monoclonal antibodies (Abcam, catalog # ab11826) were then added at a 1:800 dilution (diluted in blocking solution) at room temperature for one hour. Cells were washed three times with blocking solution, 10 minutes each. Cells were then incubated with a 1:100 dilution of goat-anti mouse Alexa Fluor 488 and Alexa Fluor 568 conjugate (Molecular Probes) containing a 1:2000 dilution of the nucleic acid stain DraQ5 (Biostatus, Shepshed, UK) for one hour at room temperature. Cells were washed with blocking solution three times, 10 minutes each. Finally, cells were washed once with 1X PBS, the chambers removed and the cells mounted and viewed on a Zeiss 510 LSM confocal microscope with an Argon/Krypton laser (Zeiss, Gottingen, Germany).

2.2 Cell culture of cell lines

The CHO (Chinese hamster ovary) cells (Kao 1968), mouse lens-derived αTN4-1 cells (Yamada, Nakamura et al. 1990), human embryonic kidney 293T cells and the human liver-derived HepG2 cell line (ATCC, Manassas, VA) were cultured in Dulbecoo's Modified essential medium (DMEM) supplemented with 10% fetal bovine serum, 1% penn-strep and 1% L-Glutamine at 37°C and 5% CO_2.

2.3 Yeast two-hybrid assay

2.3.1 Preparing the constructs

The constructs of human Prox1 homeo-Prospero domain (HDPD, amino acid 547 to 737), homeodomain only gene (HD, amino acid 547 to 642) and Prospero domain only gene (PD, amino acid 636 to 737) in pGBKT7 vector were a gift from Xiaoren Chen (University of Delaware). These were cloned by PCR using the pGAD10-human Prox1 plasmid (Cui, Tomarev et al. 2004) as template and inserted into pGBKT7 vector (Clontech, CA). The primers for HDPD were: 5' CCC GAA TTC GCC GAA GGG CTC TCC TTG TCG CTC ATA AAA 3' and 5' AAC TTA CAT ACT TCT CAT CGT CAG GAG CCT AGG GGG 3'. The reverse primer for HD was 5' -TAA GTC TAC CTC TTC ATG CGT GCA GTT CGG ATT CCT AGG GGG – 3'; the forward primer for PD was 5' - CCC GAA TTC GAG AAG TAC GCA CGT CAA GCC ATC - 3'.

SC35 construct was prepared by amplifying an in-frame region of sc35 mRNA corresponding to the C-terminal sc35 RS domain via RT-PCR using primers flanked by EcoRI and BamHI restriction sites. The forward primer for SC35 was 5'-GGG GAA TTC AGC CGC AGC CCT CGG CGA CGC -3'; the reverse primer was 5' – GGG GGA TCC GCA TTC ATC ATT TTC TTA GGA AGA – 3'. All RT-PCR reactions were done using the one-step RT-PCR kit from Qiagen (Valencia, CA). Briefly, the reaction contained 10 μl 5X buffer, 2 μl of dNTPs (10 mM each), 2 μl of enzyme mix, 1 μg of template RNA, and forward and reverse primers to a final concentration of 0.6 μM. Sufficient ddH_2O was added to bring the total volume to 50 μl. Two negative controls were prepared including one without reverse transcriptase and one without the RNA template. The cycle program was 30 minutes at 50°C, 15 minutes at 95°C, then 25-35 cycles including 1 minute at 94°C, 1 minute at (Tm-5) °C and 1 minute at 72°C. The final extension was 10 minutes at 72°C. The amplified region was then ligated into the bacterial cloning vector pPCR Script (Strategene) and transformed into competent DH5a-T1 cells (Invitrogen, CA) via electroporation.

Plasmid rescue via Miniprep yielded large quantity of bacterial cloning vectors containing the sc35 fragment. DNA sequencing then confirmed an in-frame insertion of sc35 fragment. The sc35 fragment was then cut from the bacterial cloning vector and ligated into Matchmaker pGADT7 Gal-4 AD vector. Similarly, Homeo and Prospero domains of Prox1 were previously ligated into pGBKT7 Gal4 BD

vector. This resulted in the sc35 RS domain fused to Gal4 activation domain and Homeo/Prospero domain of Prox1 fused to Gal4 binding domain.

2.3.2 Yeast transformation

Yeast strains were transformed following the YEASTMAKER yeast transformation system 2 user manual (Clontech). First, frozen stock of AH109 and Y187 haploid yeast strains were streaked on two YPDA agar plates and let grow at 30 °C for 3-4 days. One large fresh colony of each strain was inoculated into 3 ml of YPDA medium and let grow with shaking for 8 hr. These cells were then concentrated by transferring 5 μl of the culture to a 250ml sterile flask containing 50 ml YPDA and incubated at 30 °C with shaking at 250 rpm for 16-20 hr. The cultures were spun down at 700g for 5 minutes and the cell pellet resuspended in 100 ml of YPDA. Incubation at 30 °C for another 4 hr until the OD_{600} reached 0.4-0.5 followed. The yeast cells are now concentrated enough for transformation. But first, they need to be made competent. This is done by first spinning down the cells and washing the cell pellet with sterile water and then resuspending the pellet in 3 ml of 1.1X TE/LiAc solution. Next, the cells are split into two 1.5ml microcentrifuge tubes and centrifuged at 16000rpm for 15 seconds. After discarding the supernatant, the cell pellet is again resuspended in 600μl 1.1X TE/LiAc solution. The cells are now competent.

Transformation is done by adding 0.5 μg plasmid DNA, 5 μl denatured Herring testis carrier DNA and 50 μl competent cells (prepared above) in a sterile

prechilled 1.5ml microcentrifuge tube. To this tube is also added .5ml PEG/LiAc solution and the solution is incubated at 30 °C for 30 minutes with periodic mixing every 10 minutes. After 30 minutes of incubation, 20 µl DMSO is added and the solution is again incubated at 42 °C for 15 minutes with vortexing every 5 minutes. The cells are spun down in a microcentrifuge for 15 seconds, the pellet resuspended in 1 ml YPD Plus Liquid Medium and incubated at 30 °C with shaking for 90 minutes. The cells are spun down for the last time and pellet resuspended in 1 ml of 0.9 % NaCl solution. The mixture is then plated on SD/-Trp or SD/-Leu agar plates and incubated at 30 °C until colonies grow.

2.3.3 Plasmid rescue from transformed yeast cells

Before mating the two yeast strains for two-hybrid experiment, we first ensure that they have the correct plasmid insert by sequencing the plasmid DNA. To rescue plasmids, the Y-DER yeast DNA extraction reagent kit (Pierce) was used. A positive colony was allowed to grow in 10ml high stringency medium overnight. The cells were centrifuged and the cell pellet transferred to 1.5 ml microcentrifuge tube where they were resuspended in 20 µl of the Y-PER reagent. Homogeneity was ensured by pippetting the mixture up and down several times. The cells were then incubated at 65°C for 10 minutes, followed by centrifugation at 13000g for 5 minutes. The supernatant was discarded and to the cell pellet was added 16 µl of DNA releasing reagent A, and 16 µl of DNA releasing reagent B. This solution was

incubated at 65°C for 10 minutes, followed by addition of 8 µl of protein removal reagent. The mixture was centrifuged again at 13000 g for 5 minutes and the supernatant transferred to another 1.5 ml tube. The genomic DNA was precipitated by addition of 25 µl isopropanol and centrifuging at 13000g for 10 minutes. The resulting DNA was finally resuspended in 50 µl sterile water.

In order to sequence the rescued plasmid, it was first transformed in *E. coli* cells. To transform electrocompetent cells (Stratagene, CA), Electroten-blue electroporation competent cells were thawed on ice and 40 µl of cells aliquoted into chilled 1.5 ml microcentrifuge tube. Ten µl of the rescued DNA was added and contents were mixed gently. This DNA-cell mixture was transferred to a chilled BTX electroporation cuvette with 0.1 cm gap (Genetronics, CA). The cuvette was then placed into a chilled electroporation chamber. Transformation was achieved by electroshocking the cells using a pulse of 1700V electricity. 960 µl of sterile SOC medium (Invitrogen) was immediately added to resuspend the cells. The entire mixture was transferred to a 14ml BD Falcon tube and incubated at 37°C for 90 minutes with shaking at 250 rpm. Finally 10 µl and 100 µl of the transformation mixture were spread on LB-Ampicillin agar plates respectively. The plates were incubated at 37 °C overnight.

For each transformation, 4 colonies were chosen for culturing overnight in 5 ml LB medium and plasmids were subsequently rescued by Miniprep (Promega, WI). The DNA was then digested with EcoRI/BamHI and subjected to 0.8% agarose

gel. After confirming that all the colonies have the same size of insert on the gel, one of the plasmids was sent for sequencing.

2.3.4 Yeast mating

Yeast strain AH109 containing the bait pGBKT7-Prox1 HDPD and Y187 containing the prey pGADT7-SC35 constructs were prepared as described in 2.3.2. These two strains were mated following the Clontech protocol. Briefly, one large fresh colony of AH109 containing pGBKT7-Prox1 was inoculated in 50ml SD/-Trp medium and incubated at 30 °C overnight with shaking at 250 rpm. The cells were spun down at 1000g for 5 minutes and resuspended in the residual medium (around 5ml). Same was done for the Y187 strain containing pGADT7-SC35 vector. One ml of each strain was placed in 2L sterile flask and 48ml of 2X YPDA/Kan medium added to bring the total volume to 50 ml. This mixture was then incubated at 30 °C overnight with shaking at 250 rpm. After 24 hr of mating, the cells were spun down by centrifuging at 1000g for 10 minutes and the pellet resuspended in 10ml of 0.5X YPDA/Kan medium.

150 µl of the mating mixture was plated on SD-Trp/-Leu double dropout low stringency media containing X-α-Gal. X-α-Gal was added to the plate for detection of the MEL reporter gene product, α-galactosidase. The plates were incubated at 30 °C until colonies appear (usually 5-6 days). The positive colonies should be robust, blue and large. Only the cells that contain both pGBKT7-Prox1 and

pGADT7-SC35 will grow on this medium. The positive colonies were then restreaked on high stringency, quadruple dropout SD-Trp/-Leu/-His/-Ade media containing X-α-Gal. Colonies will only grow if Prox-1 and sc35 interact. For positive control, we used pGbKT7-p53 and pGADT7-T antigen.

2.4 CAT assay

Cells were transfected following the Lipofectamine-Plus reagent protocol provided by Invitrogen. Accordingly, $3.6*10^5$ cells were plated for each 60-mm dish the day before transfection. The plasmids added included 0.25 μg of pCMV/β-GAL, 3 μg of promoter/ chloramphenicol acetyltransferase (CAT) plasmid and 0.25 μg to 1 μg of expression plasmid. Cells were harvested 48h after transfection and cellular extracts were prepared by multiple cycles of freeze/thaw. The extracts were assayed for CAT as described below.

5 μl of cell (15 μl PLE cell) extract diluted by 15 ul lysis buffer was heated at 65 °C for 15 minutes. The extract from each sample was then mixed with 25 μl of 1 M Tris pH 7.8, 50 μl of 5 mM chloramphenicol, 6 μl of ^{3}H-acetyl CoA (Perkin Elmer, 0.25 mCi/500 μl), and 119 μl H_2O. 200 μl of this mixture was pipetted under 5 ml of Packard Econoflor (Perkin Elmer) in a scintillation vial. The counts per minute (cpm) for each sample were measured seven times, at approximately 15 minute intervals in a Packard BioScience Tri-Carb 2900TR Liquid Scintillation Analyzer using Quanta Smart 1.31 software (Perkin Elmer). The CAT activity was defined as

cpm/min, which corresponded to the rate of production of ^{3}H- acetylated chloramphenicol. The cpm/min for each sample was normalized against β-gal activity as a control for efficiency of transfection.

2.5 Prox1 structure prediction

Molecular modeling of Prox1 structure was carried out on a Silicon Graphics O2 workstation, using the programs InsightII, and Discover, as well as the Swiss-Model online workbench.

2.5.1 Homology modeling

The X-ray solved Prospero structure available at the PDB (PDB accession code 1MIJ, and 1XPX), share 66.25% sequence identity with Prox1 in the homeo-Prospero domain (HDPD) regions near the C-terminus and thus could be used as a template to build rather accurate three-dimensional model for Prox1 HDPD. Homology model prediction was carried out using the Swiss-Model server (Guex 1997; Schwede T 2003; Schwede 2004).

2.5.2 Structure refinement

Steric conflicts were corrected using the rotamer library and the search algorithm implemented in the Homology program of InsightII (Richards 1987; M.T.

Mas 1992). The models were then subjected to an extensive minimization using a combination of the steepest descent and conjugated derivatives algorithm with an energy tolerance of 0.001 Kcal/mol. Energy minimizations were done using the cvff forcefield of Discover.

2.5.3 Structure Validation

The predicted Prox1 HDPD model was validated by inspection of the geometrical parameters from the Ramachandran plot, created using the program RAMPAGE (S.C. Lovell 2002). The program Superpose was used to perform a superposition of the predicted model with the template protein (Rajarshi Maiti 2004) and PROCHECK (R.A. Laskowski 1993) was used to assess the geometric quality of the three-dimensional models. As an example, 79% of the residues were correctly assigned on the best allowed regions of the Ramachandran plot, the remaining residues being located in the generously allowed regions of the plot except for Glu660 which occurs in the non allowed region.

2.6 Docking Analysis

The structure of Prox1 was docked with different fragments of the protein PCNA (pdb accession codes 1PLR, 1AXC) and short fragments of DNA molecule. In all cases, docking of receptor (Prox1) and ligand (either PCNA or DNA) was carried out using ZDOCK for the initial-stage docking which optimizes for desolvation, grid-

based shape complementarity and electrostatics to generate multiple plausible protein-ligand poses (Chen R 2002; Chen R 2003; Chen R 2003). Ranking and refinement were then carried out using RDOCK which energy minimizes the predicted complexes using CHARMM (A. D. MacKerell 1998) force field. Subsequently, each structure was ranked based on the desolvation and electrostatic energy (Li L 2003).

The three most favorable complexes were subjected to a short molecular dynamics (MD) calculation. All MD calculations were performed using the Discover module. The minimized model was heated from 0^0 K up to 300^0 K for 9 ps with a 5^0 K increment every 100 time steps (1 time step = 0.0015 ps). Once at 300^0 K, the model was subjected to a 150 ps simulation. An average structure of the three most stable structures encountered during the simulation was calculated and further minimized with an energy tolerance of 0.001 Kcal/mol. This is now the final configuration of receptor ligand complex.

The solvent accessible surface area of the complexes was also computed using NACCESS (Hubbard 1993).

2.6.1 Prox1 PCNA docking models

PCNA structures were retrieved from the PDB (PDB accession codes 1PLR, 1AXC). Three structural fragments of the PCNA molecule were used to dock with Prox1: full length trimer PCNA, monomeric PCNA, and only C-terminal region

(amino acids 167 to 261). During the docking stage, the PIP box residues of Prox1 were forced into the binding site.

2.6.2 Prox1 DNA docking models

The B-form double stranded DNA structures of the OL2 site (5'-CACTTCCA-3'), -220 site (5'-TGC GGC AAA GTG GCG CGG-3'), and -290 site (5' AGT GCT GGA TCC AGG TGC TGG 3') were built using the Biopolymer module of Insight II program. Each structure was subjected to a short energy minimization cycle (200 steps). Docking was then performed as described above.

Chapter 3

MONOCLONAL ANTIBODIES AGAINST PROX1

In order to obtain relevant morphological data on the expression pattern of Prox1 and for use as a reagent in numerous experiments, including Western blot, and immunoprecipitation, a highly specific mouse monoclonal antibody against Prox1 was developed in our lab.

Two monoclonal Prox1 antibodies, 5G10 and 4G10, with different isotype against the highly conserved human Prox1 homeo and Prospero domain were developed in our lab; for procedural details, refer to (Chen X 2006). Here we report usability of these antibodies in immunofluorescence of tissue sections. Ascites and purified antibodies derived from hybridomas were used to detect Prox1 in vertebrate lenses, a tissue known to express Prox1 at appreciable levels (Duncan, Cui et al. 2002). Hybridoma 5G10 was able to robustly detect nuclear Prox1 protein in sections derived from human, (Figure 11 A-C), rat (Figure 11 D-F), chicken (Figure 11 G-I), and lizard (*Anolis sagrei;* Figure 11 J-L) lenses. Frog (*Rana pipens*) lenses only stained weakly with this antibody while fish (*Fundulus similes*) lenses did not stain at all (data not shown). Hybridoma 4G10 generally produced the same pattern of results

(data not shown), however, 4G10 generally stained tissue less intensely than 5G10 when used at the same dilution.

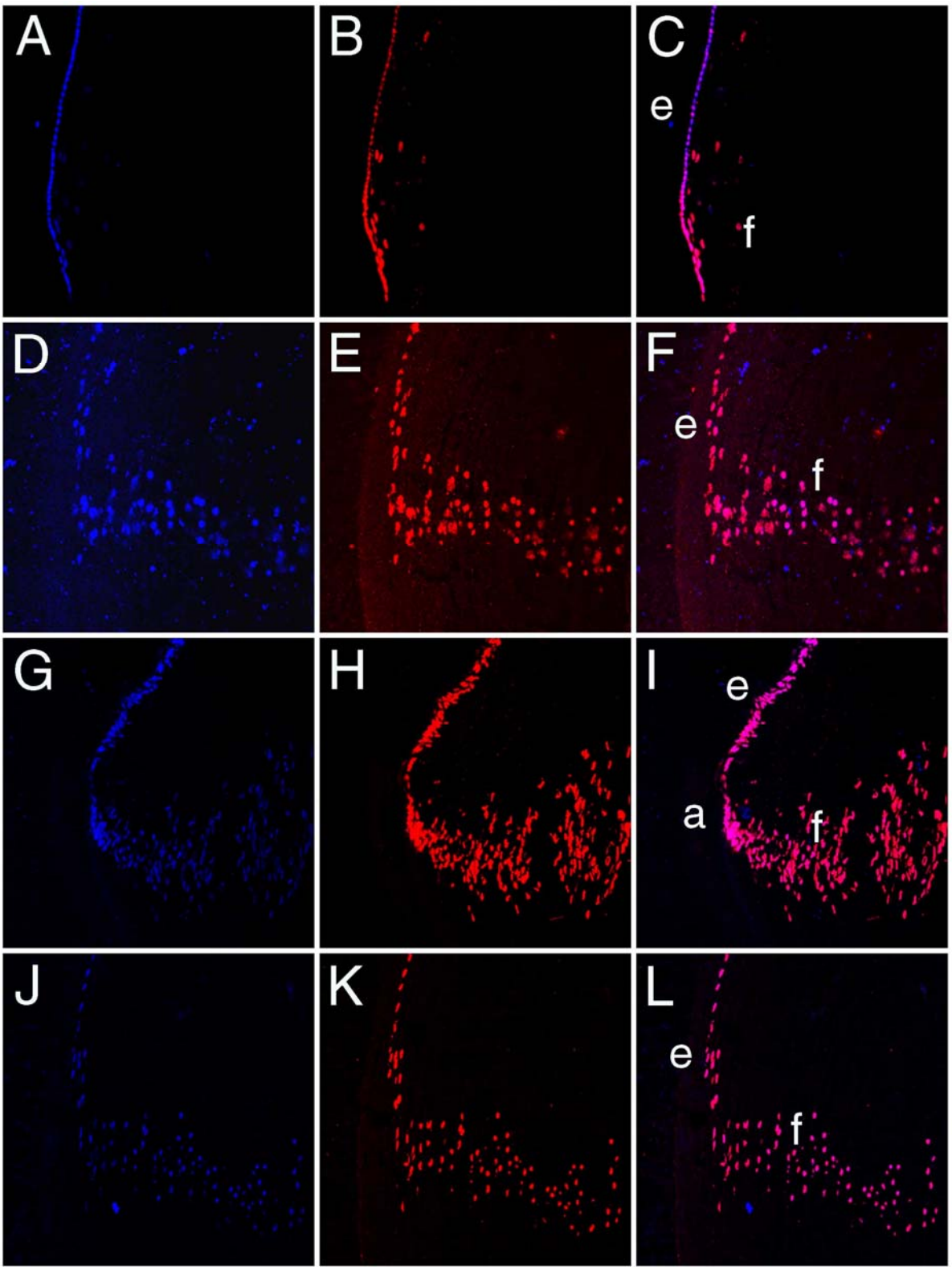

A
B
C
e
f
D
E
F
e
f
G
H
I
e
a
f
J
K
L
e
f

Figure 11: Immunofluoresence staining of lenses from various terrestrial vertebrates with the 5G10 monoclonal antibody (1:2000 dilution). A-C 21-year-old human lens. D-F Five month rat lens. G-I Embryonic day ten *Gallus gallus* (chicken) lens. J-L Adult *Anolis sagrei* (lizard) lens. a, annular pad; e, lens epithelium; f, lens fiber cells. A, D, G, J- the nuclear counterstain TO-PRO-3 in blue; B, E, H,K- Prox1 immunodetection in red; C, F, I, L- Overlap in pink.

Both antibodies detected Prox1 in lenses from all land vertebrates tested from humans to lizards by immunolocalization. Further, they can detect both endogenous and eukaryotic Prox1 expression vector-driven Prox1 expression in cell and tissue extracts by western blot (data not shown) demonstrating that they are robust reagents with diverse potential applications in the study of Prox1 function in tissue development and disease.

Prox1 monoclonal antibodies may actually serve a much wider purpose. Immunostaining of lymphatic tissues from healthy human adults and lymphedema patients with Prox1 polyclonal antibodies showed that Prox1 is a reliable and highly specific marker for lymphatic endothelial cells in normal and pathologic human tissues (Wilting, Papoutsi et al. 2002; Reis-Filho and Schmitt 2003; Van der Auwera, Van Laere et al. 2004; Al-Rawi, Mansel et al. 2005). However, the routine use of Prox1 staining in the clinic to identify lymphatics in biopsy specimens is impeded by the lack of highly standardized and reproducible anti-Prox1 monoclonal antibodies. Thus, the highly specific monoclonal antibodies are well suited for use in the clinic as well as in basic science laboratory setting.

Chapter 4

PROX1 SC35 INTERACTION

4.1 Library scale yeast two-hybrid

As noted earlier, Prox1 is not only a transcription factor, in the classical sense, but can also bind to other transcription factors and modulate their function. This led us to believe that perhaps the function of Prox1 can also be modulated by its interaction with other proteins. If so, it would help explain how a single transcription factor can influence such diverse gene control.

In order to test this hypothesis, the interacting partners of Prox1 were first identified using a library scale yeast two-hybrid assay carried out by Xiaoren Chen in our lab. The yeast two-hybrid uses a reporter gene to detect the physical interaction of a pair of proteins inside a yeast cell nucleus. This system is designed so that when a target protein binds to another protein in the cell, their interaction brings together two halves of a transcriptional activator (Gal4), which is then able to switch on the expression of the reporter gene. Library scale yeast two-hybrid screen was done using Prox1 fused to the DNA binding domain of Gal4 and a library of proteins derived from the lens of an 11-day embryo mouse cDNA fused to the Gal4 activation domain. Chen found that Prox1 interacted with various proteins including nuclear proteins,

cytoplasmic proteins and membrane proteins. Although most interacting proteins are involved with cell cycle control (as expected, since Prox1 plays a major role in cell fate determination), there were also multiple positive hits for interferon response element-binding factor (IREBF-2)/ Splicing factor 35 (SC35). This suggested a *bona fide* interaction between Prox1 and SC35 and so I chose to study this interaction in more detail.

4.2 SC35

SC35 is a well studied protein. It is a member of serine and arginine (SR) rich class splicing factors that have dual roles in pre-mRNA splicing. The SR proteins are closely related, highly conserved RNA-binding proteins that are essential for constitutive splicing of the pre-mRNA and also regulate splicing in a concentration-dependent manner by influencing the selection of alternative splice sites (Ge 1990; Krainer 1990; Krainer 1990; Zahler 1993). In humans, the SR protein family is encoded by nine *splicing factor*, *arginine/serine-rich* genes, designated *SFRS1-7*, *SFRS9*, and *SFRS11*.

Table 1: Human genes encoding SR proteins (Brenton R Graveley 2005)

Gene	SR protein
SFRS1	SF2/ASF/SRp30a
SFRS2	SC35/SRp30b
SFRS3	SRp20
SFRS4	SRp75
SFRS5	SRp40
SFRS6	SRp55
SFRS7	9G8
SFRS9	SRp30c
SFRS11	SRp54

All members of the SR family have a common structural organization with two modular domains, an N-terminal RNA recognition motif (RRM) and a C-terminal domain of variable length rich in serine and arginine residues that interacts with other proteins involved in splicing reactions. Members of the SR family have either one or two RRM motifs. SC35 specifically contains only one RNA recognition motif. The reader may be more familiar with the SR proteins as belonging to the nuclear subregions called speckles.

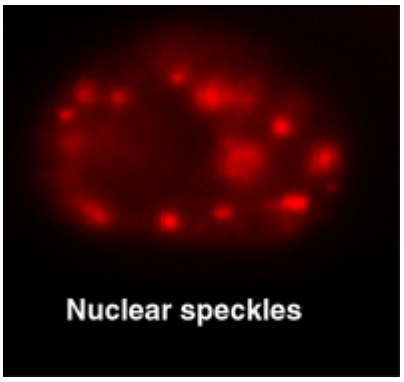


Figure 12: Staining of cells with sc35 antibodies shows distinct clusters of localized protein, called nuclear speckles.

Speckles are subnuclear structures that are enriched in pre-messenger RNA splicing factors, like sc35, and are located in the interchromatin regions of the nucleoplasm of mammalian cells. The speckle domain seen by immunofluorescence corresponds to interchromatin granule clusters(IGCs) and perichromatin fibrils at the electron microscope level (Spector 1993). As their name suggests, IGCs are distributed in regions of the nucleoplasm that seem to be depleted of chromatin (Thiry 1995). Speckles typically appear as irregular, punctate structures and can vary in size and shape. The nuclear organization of splicing factors is dynamic, and it has been proposed that SR proteins and other splicing factors are recruited from the interchromatin granule clusters, which are believed to be sites of storage or assembly, to the sites of active transcription (Jiménez-García 1993; Huang 1996; Misteli 1997; Spector 2003). Consistent with this view is the experimental observation that when transcription is halted, with either transcriptional inhibitors or heat shock, splicing speckles become rounded and larger, and the splicing factors redistribute from sites of transcription back into the splicing speckles (Mintz 1999). Although not directly

involved in transcription, the splicing factors are needed to recruit the U-snRNPs (U1,U2,U4-6) to the correct splice sites, as illustrated in the figure below.

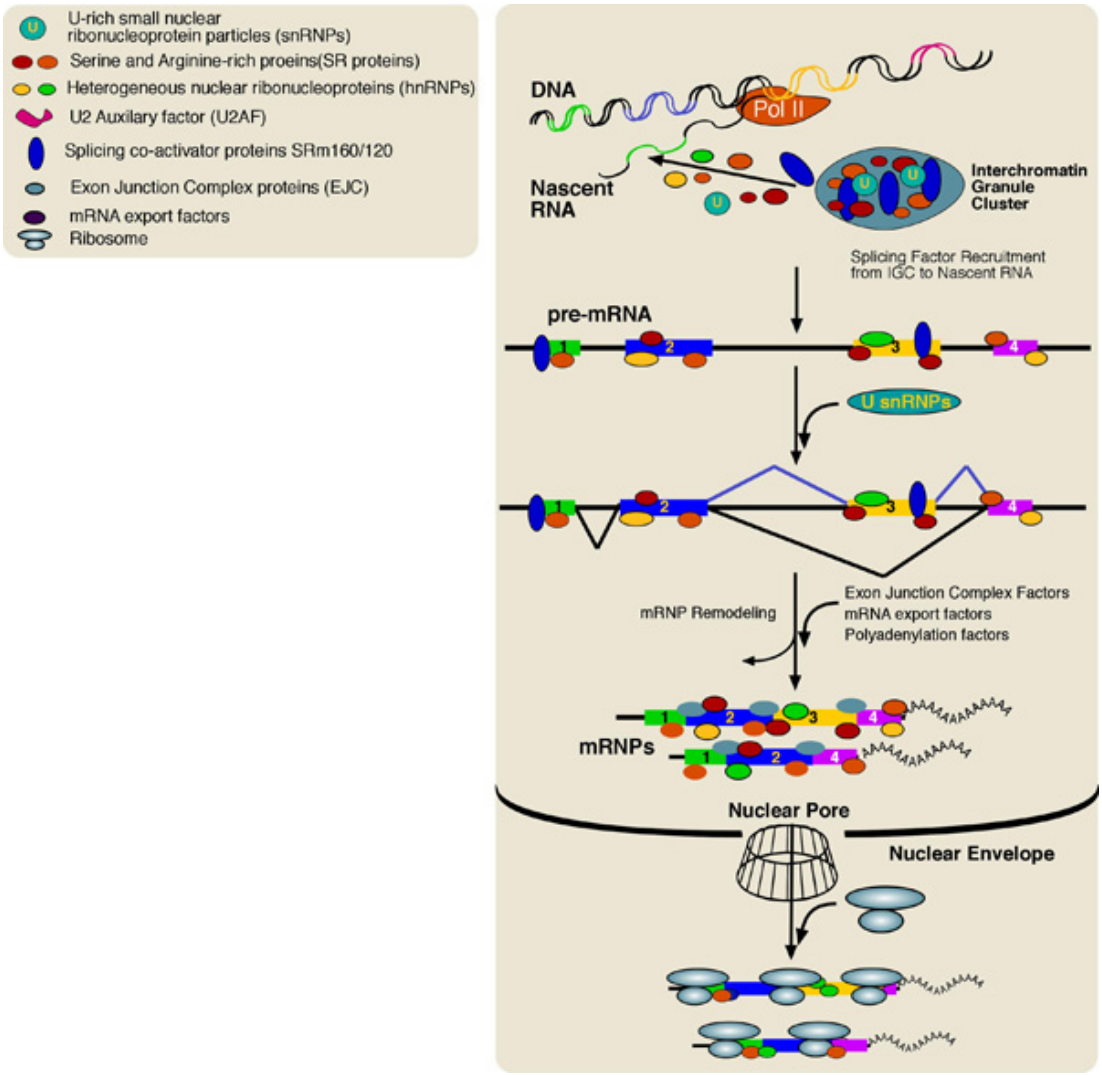


Figure 13: Steps in RNA processing. As a nascent pre-mRNA is synthesized, splicing factors (localized to the interchromatin granule cluster) are recruited to the pre-mRNA and bind to specific sites through their N-terminal RNA recognition motifs. Binding of splicing factors helps recruit the U-snRNPs to the correct splice sites, thus forming a spliceosome assembly. Splicing of introns through an intermediate lariat structure follows.

More specifically, there are three proposed models for the role of sc35 in splicing. In one model, the U1 snRNP recruitment model, sc35 first binds to an upstream exon and recruits U1 snRNP at the 5' splice site. Alternatively, enhancer bound sc35 aids in recruiting the U2 auxiliary factor (U2AF), which subsequently recruits U2 snRNP to the branch point A. In both these models, the splicing factor physically binds to the pre-mRNA. The last model proposes an exon-independent function of SR proteins in which they facilitate the splice site pairing by (1) simultaneously interacting with U1 snRNP and U2AF across the intron, and (2) by recruiting the U4/U6/U5 tri-snRNP. These models are depicted below; sc35 is replaced with a generic label, SR protein.

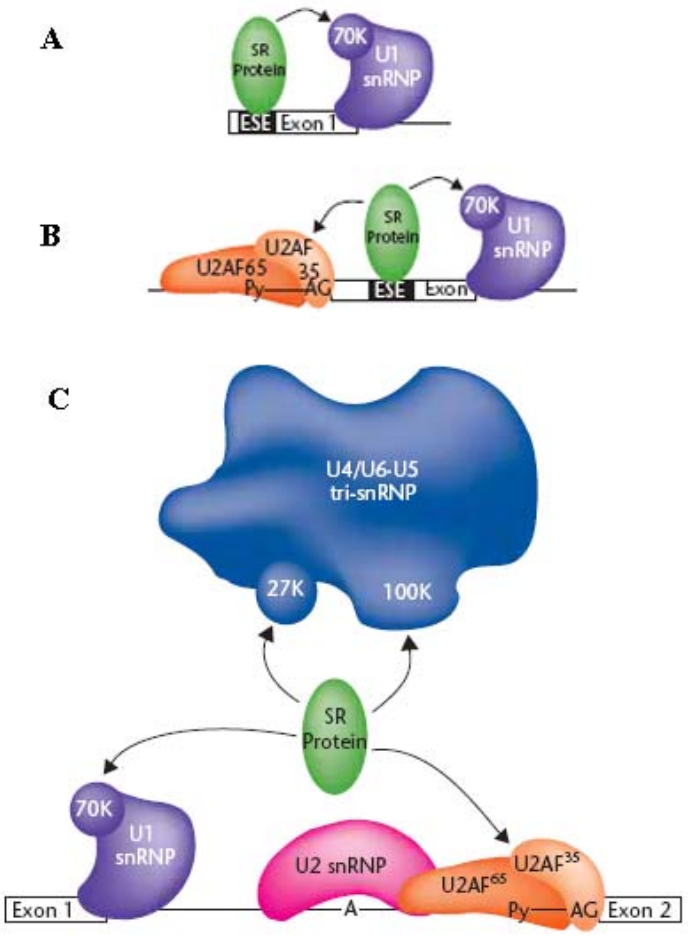


Figure 14: Exon-dependent (A,B) and exon-independent (C) functions of SR proteins. A: Exon bound SR protein recruits the U1 snRNP to the 5' splice site. B: Enhancer bound SR protein recruits U2AF35 which then recruits the U2 snRNP to the branch point A (not shown). C: In the exon-independent model, the SR protein simultaneously interacts with both U1 and U2AF snRNPs. Figure modified from (Brenton R Graveley 2005).

4.3 Confirming the Prox1-SC35 interaction

When the library size yeast two-hybrid screen was performed by Chen, sc35 and several other positive hits were in an incorrect reading frame with respect to the Gal4 activation domain. However, this is not problematic since yeasts are known to skip reading frames and still be able to synthesize the correct, in-frame, protein.

Before proceeding any further, it became necessary to demonstrate whether an in-frame sc35 gene is able to interact with Prox1. Since the sc35 fragment obtained from yeast two-hybrid corresponded to the C-terminus RS domain, an in-frame constructs of the C-terminus RS domain of sc35 fused to Gal4 activation domain and Prox1 HDPD fused to Gal4 DNA binding domain were created as described in chapter 2. These constructs were sequenced to ensure the presence of correct reading frame. Yeast two-hybrid results showed that these two proteins indeed interact. The interaction is also present when full length sc35 is used; however, using only the Homeo or only the Prospero domains of Prox1 diminishes interaction with sc35. Thus, we conclude that Prox1 and sc35 interact via the homeo-Prospero domain of Prox1 and the RS domain of sc35.

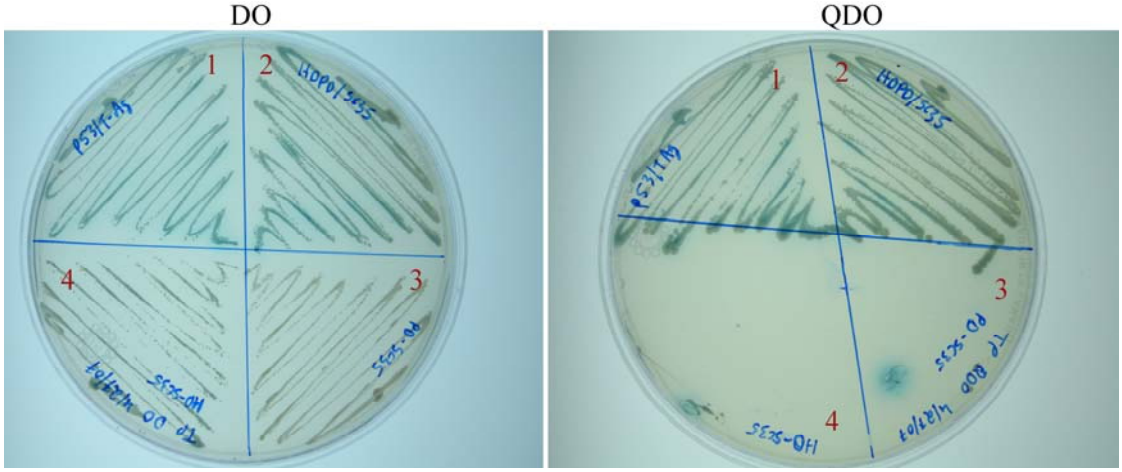


Figure 15: SC35 interacted with the Prox1HDPD, but not the HD or PD alone. Transformed AH109 yeast containing Prox1 HDPD, HD and PD were mated with transformed Y187 containing SC35 and plated on low stringency (DO) medium respectively. The growing colonies were streaked to both DO and high stringency (QDO) medium. HDPD interacted with SC35 and grew on both plates (2). But PD (3) or HD (4) alone only survived on DO plates not QDO plates. p53-T Antigen interaction was used as a positive control (1).

We give further support that Prox1 and sc35 interact by establishing a colocalization pattern in the lens tissue, where both proteins are known to be present. Colocalization was done using mouse monoclonal sc35 antibodies and rabbit polyclonal Prox1 antibodies in E47 day human lens vesicle.

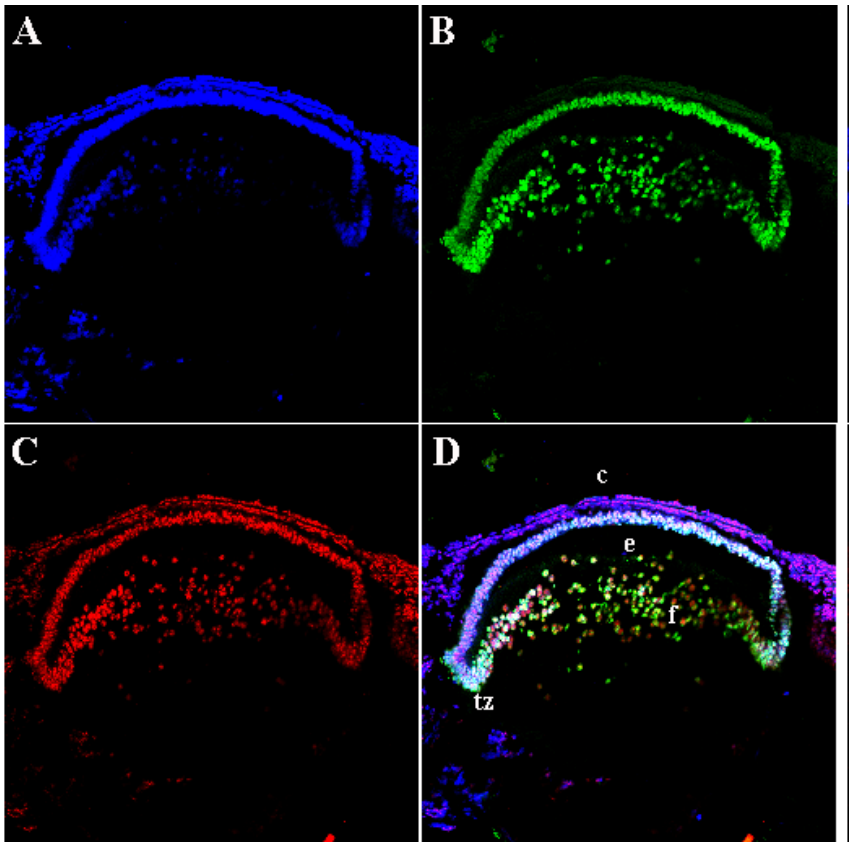


Figure 16: Prox1 and SC35 are strictly colocalized in E47 day human lens vesicle. A: Nuclear staining in blue; B: SC35 in green; C: Prox1 in red; D: overlap. c-cornea; e-lens epithelial cells; f-lens fiber cells; tz-transition zone.

Colocalization data, along with the established interaction in yeast, give strong support that Prox1 and sc35 interact *in vivo*.

4.4 Consequence of Prox1 SC35 interaction

An interaction between a splicing factor and a transcription factor is rather unusual. Transcription and splicing are removed processes; however, they both have the same ultimate goal – to synthesize an mRNA molecule. Thus, it seems plausible that there be a link between splicing and transcription and the interaction between Prox1 and sc35 may provide an example of such a link. Thus, in order to fully describe the paired interaction, we investigate the consequence of sc35 interaction on Prox1 function.

4.4.1 SC35 represses Prox1 activation to βB1-crystallin promoter

As noted previously, Prox1 plays many different roles in different tissues. However, in lens fiber cells, we know that Prox1 acts as a true transcription factor and binds to the OL2 element of the chicken βB1-crystallin promoter thereby activating the promoter for transcription (Wenwu Cui 2004). Thus, we decided to test whether the Prox1-sc35 interaction can affect the activation of the βB1-crystallin promoter from the OL2/PL2 element. This experiment was done by co-transfecting full length sc35 and Prox1 into Chinese hamster ovary (CHO) cells along with a reporter vector consisting of three consensus Prox1-responsive OL2/PL2 elements derived from the chicken βB1-crystallin promoter placed upstream of the CAT reporter gene that is driven by the βActin basal promoter.

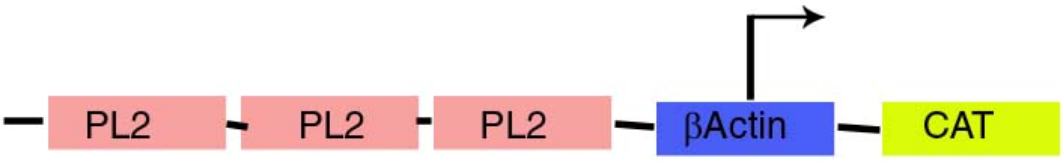


Figure 17: Plasmid construct used for CAT assay. Three copies of the OL2/PL2 element are placed upstream a minimal βActin promoter from which the transcription of *CAT* initiates.

CAT assay using only Prox1 strongly transactivated the 3XOL2/PL2 element. When Prox1 and sc35 were co-transfected, sc35 significantly repressed Prox1 activation to the OL2/PL2 element ($P<0.01$). The control experiments showed that the repression was specific to the OL2/PL2 element, and not to the βActin basal promoter.

Furthermore, we found that the repression is a concentration dependent phenomenon. High levels of sc35 (1 μg), while keeping Prox1 levels constant (.5 μg), almost abolished Prox1 activation, and low levels of sc35 (.05 μg) had very little effect on Prox1 activation to the OL2/PL2 element. Similarly, over-expression of Prox1 (1 μg), while keeping sc35 levels constant (.5 μg), resulted in slightly larger repression than when moderate amount of Prox1 (.5 μg) was used.

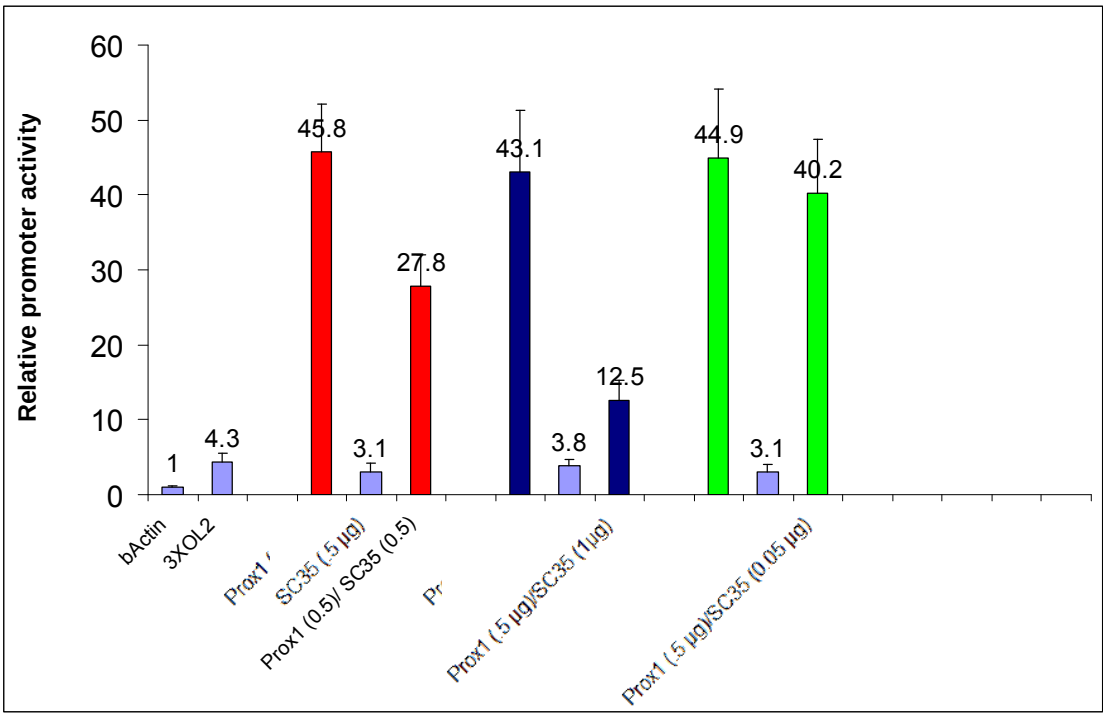


Figure 18: Prox1 transactivation to the OL2 element of βB1-crystallin promoter was repressed by SC35 in a dose-dependent manner. Prox1 (0.5 µg) activated the 3XOL2 element upstream of the βActin minimal promoter and the activation was repressed by SC35 (.5 µg). This repression was specific to the OL2 element but not the βActin minimal promoter. Increasing the levels of SC35 while maintaining a constant Prox1 concentration almost completely abolished Prox1 activation.

The data from CAT assays using different concentrations of Prox1 and sc35 are summarized in table below. Here we report the percent repression of the 3XOL2 element upon addition of sc35. The values are simply computed as, (relative reporter activity using Prox1 alone – relative reporter activity with both Prox1 and sc35)/(relative reporter activity using Prox1 alone).

Table 2: Percent repression of Prox1 upon addition of variable amounts of sc35 and Prox1.

	Prox1, 0.5 µg	Prox1, 1.0 µg
SC35, 0.05 µg	12.2%	7.1%
SC35, 0.5 µg	38.0%	61.5%
SC35, 1.0 µg	72.7%	84.5%

4.4.2 Intracellular distribution of sc35

Since SC35 is most readily observed by fluorescent antibodies, we investigate whether the Prox1 sc35 interaction affects the distribution of sc35 in cultured cells. To this end, we co-stained cultured αTN4-1 cells with sc35 and Prox1 antibodies. The αTN4-1 cell line is derived from mouse lens fiber cells and natively expresses both Prox1 and sc35. Fig 18 shows that Prox1 protein expression is variable among cells, but those cells producing large amounts of Prox1 (bright red staining in panel C), show a corresponding redistribution of the nuclear speckles, indicated by the arrows. SC35 staining shows the characteristic nuclear speckles with staining localized to multiple clustered regions in the nucleus; however, these speckles become diffuse in cells that also have large amounts of Prox1; furthermore, the sc35 distribution in these cells overlaps better with the Prox1 distribution. This leads us to believe that the Prox1 sc35 interaction causes the export of sc35 from the IGCs.

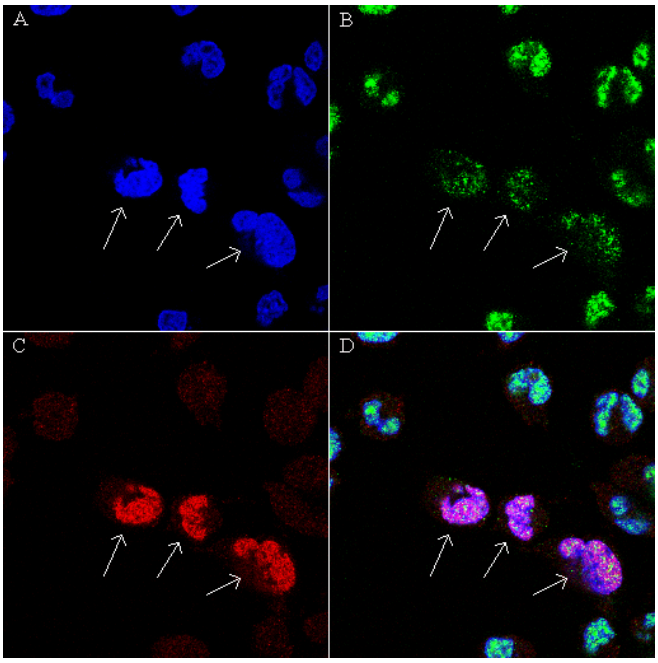


Figure 19: Co-staining of cultured αTN4-1 cells with Prox1 and sc35. A: nuclear stain in blue; B: sc35 staining in green; C: Prox1 staining in red; D: overlap. Arrows indicate cells that have undergone speckle redistribution.

In order to test whether Prox1 changes the distribution of sc35 protein in the nucleus, we transfected full length Prox1 into CHO cells and observe the resulting sc35 distribution by immunofluorescence. CHO cells do not natively express Prox1, but do contain sc35, thus the introduction of exogenous Prox1 should change the sc35 distribution. This experiment, however, did not yield expected results. Transfection of Prox1 in CHO cells did not alter sc35 distribution. These conflicting observations are remedied in our discussion.

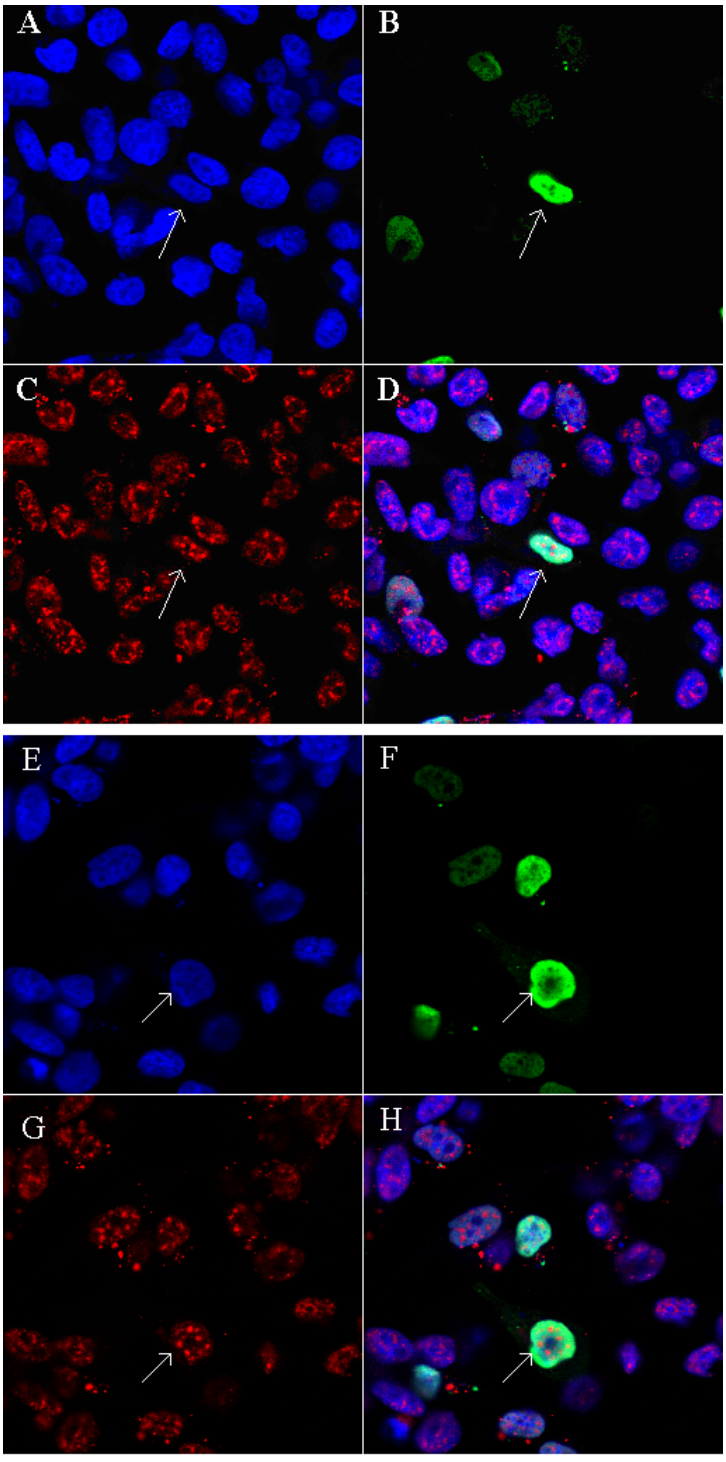


Figure 20: Two panels showing the lack of redistribution of sc35 in Prox1 transfected CHO cells (arrows). A,E: nuclear stain in blue; B,F: Prox1 in green; C,G: sc35 in red; D,H: overlap.

4.5 Discussion: Role of sc35 in Prox1 function regulation

By a series of small scale yeast two-hybrid experiments, we determined that the C-terminal RS domain of sc35 interacts specifically with the homeo-Prospero domain of Prox1. The consequence of this interaction on Prox1 function was tested through a CAT assay in which we found that sc35 represses Prox1 mediated activation of the chicken βB1-crystallin promoter through the OL2/PL2 element. Furthermore, the repression of Prox1 function is dependent on intracellular sc35 concentration. Also, we report that sc35 and Prox1 are both strictly colocalized in the developing human lens vesicle, and remain colocalized in the lens epithelial and lens fiber cells of adult human (data not shown). However, when cultured αTN4-1 cells are viewed at higher magnification, we find that although Prox1 and sc35 are present in the same cell, their staining patterns do not overlap. Rather, we consistently observed a more diffused sc35 staining in cells that also contained high amounts of Prox1, and a clustered sc35 staining in cells with lower amounts Prox1. Similar observations are made in HepG2 cell lines, indicating that this is not an isolated occurrence. However, when CHO cells, which do not natively express Prox1, are transfected with Prox1, the expected change in sc35 distribution is not observed.

We can rationalize these finding with the following model. First note that the nuclear speckles (site of sc35 storage) are dynamic structures. During high transcriptional activity, the splicing factors are recruited to the site of pre-mRNA synthesis where they participate in splicing reactions. Thus, the speckles become

diffused throughout the nucleoplasm. However, during low transcriptional activity, majority of the splicing factors are not actively involved in splicing (because there is not much pre-mRNA to splice to begin with), and they reform into tight round clusters, characteristic of classical nuclear speckles (Mintz 1999).

Thus, we propose that although sc35 and Prox1 have the ability to interact, they are compartmentalized in different regions of the nucleus and do not normally interact with one another. However, when Prox1 activates a gene for transcription, the splicing factors, including sc35, are recruited from their normal storage area (the nuclear speckles) and arrive at the site of pre-mRNA transcription. Here, the splicing factors would normally facilitate the spliceosomal assembly. If, however, the rate of pre-mRNA synthesis is high, for example due to over-expression of Prox1 as in the CAT assay, excess splicing factors would then be recruited. Here, most splicing factors bind the nascent pre-mRNA and prepare for splicing, while the excess pre-mRNA unbound sc35 is free to interact with Prox1. This interaction results in the dissociation of Prox1/OL2 complex and transcription is suppressed.

This model is consistent with the CAT assay results. Co-transfection of CHO cells with small amount of sc35 and moderate amount of Prox1 did not have significant effect on Prox1 activation of the OL2/PL2 element. However, large amount of sc35 resulted in almost abolishment of Prox1 activation. In both cases, the amount of Prox1 is kept constant and therefore, if no additional sc35 was introduced in these cells, the relative activation of the *CAT* gene from the OL2/PL2 element should be the

same. When very little sc35 is added, the small excess splicing factor is free to interact with Prox1 and pull it away from the OL2/PL2 element thus slightly repressing the reporter gene activity. On the hand, large addition of sc35 plasmid produces an excess of free sc35 protein which binds to the majority of Prox1 molecules present in the cell and interferes with the activation of *CAT* gene. Thus we observe almost complete abolishment of *CAT* activation.

Similarly, when large amount of Prox1 is transfected, and sc35 kept at a constant moderate level, we observe relatively larger repression than when smaller amount of Prox1 was transfected. In both cases, the amount of sc35 is kept constant. A large addition of Prox1 results in higher rate of *CAT* pre-mRNA synthesis, but also more sc35 is recruited to the site of DNA transcription. Once sc35 and Prox1 are localized in the same sub-compartment of the nucleus, they are able to interact. The sc35 bound Prox1 complex is not as efficient towards activating the *CAT* gene from the OL2/PL2 promoter element, thus we observe a large repression of the *CAT* activity. However, addition of smaller amount of Prox1 does not result in as great repression since the rate of mRNA synthesis is lowered and not as much sc35 is recruited from the nuclear speckles.

The immunofluorescence data showing sc35 and Prox1 distribution in cultured cells is also consistent with our proposed model. We observed that in certain αTN4-1 cells that natively express larger amounts of Prox1 have more diffused sc35 distribution, characteristic of high transcriptional activity; whereas in majority of the

cells with moderate staining for Prox1 showed a classical clustered distribution of sc35, characteristic of low transcriptional activity. Furthermore, when we transfected Prox1 into CHO cells that do not natively express Prox1, we did not observe any change in sc35 distribution. This is because even though Prox1 is present in these cells, Prox1 and sc35 are localized to different sub-compartments of the nucleus. Sc35 only leaves the interchromatin granule regions when pre-mRNA is actively being synthesized. And since Prox1 does not activate any gene transcription in CHO cells, sc35 remains localized to the nuclear speckles. The only reason we observed a change in the intracellular sc35 distribution in αTN4-1 and HepG2 cells is because here, the larger amount of Prox1 increases the rate of gene activation and pre-mRNA synthesis. In times of high transcriptional activity, sc35 diffuses out of the clustered nuclear speckles and arrives at the site of pre-mRNA synthesis.

In summary, we propose that although sc35 and Prox1 have the potential to interact, they only do so when Prox1 mediated gene activation is high. An increase in gene activation and resulting increase in pre-mRNA synthesis brings sc35 out of its localized nuclear speckle compartment and to the site of DNA transcription. Here, Prox1-sc35 interaction results in the suppression of Prox1 function as a transcription factor. Thus, sc35 is an important regulator of Prox1 function.

4.6 Future Work

In order to further validate our model of the role of sc35 in Prox1 function regulation, several experiments should be completed. First, more concrete evidence of Prox1-sc35 interaction *in vivo* can be obtained by either a pull-down assay or co-immunoprecipitation using lens tissue. The co-immunoprecipitation experiment was attempted without success, partly because either the interaction was not strong enough or because, as our model predicted, Prox1 and sc35 might be compartmentalized in different parts of the cell so the number of Prox1-sc35 interacting molecules might be too low under the experimental conditions. However, these experiments should be repeated to obtain more conclusive evidence of *in vivo* interaction.

Next, transfection of CHO cells with Prox1 did not result in a change in sc35 distribution because Prox1 does not activate any genes in CHO cells. However, if we co-transfect Prox1 and the CAT vector, containing the PL2/OL2 element, in CHO cells, immunostaining should reveal a change in intracellular sc35 distribution in cells that also contain Prox1. Furthermore, a developmental expression pattern of Prox1 and sc35 should be established. This will help identify other tissues in which sc35 can modulate Prox1 function. Also, I have focused much of my attention on the effect of Prox1-sc35 interaction on Prox1 function and have neglected to study whether the function of sc35 is at all altered. Splicing factors are known to be involved in alternative splice site recognition and it is possible that interaction with Prox1 allows sc35 to splice an alternative variant of a gene. This is an interesting possibility since that would allow Prox1 to influence gene control without actually binding to DNA.

The possibility of changing sc35 function upon Prox1 interaction should be addressed in the future by some splicing assay where one would look at the gene products spliced by sc35 in the presence and absence of Prox1.

Chapter 5

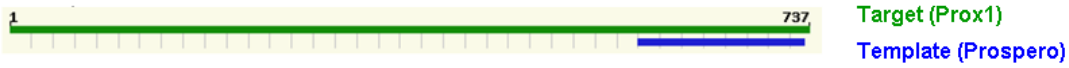
STRUCTURE PREDICTION OF PROX1

5.1 Homology model of Prox1

In order to better understand the biochemistry of Prox1 and to help explain the interactions of Prox1 with other proteins and DNA, we naturally turn to its structure. However, since the structure of Prox1 has not yet been experimentally determined, we predicted Prox1 structure using knowledge-based homology modeling technique. The knowledge based method for predicting protein structure relies on using information from large protein structure databases (such as PDB) to develop models of proteins of unknown structure. This comparative model (homology modeling) approach is most accurate when the unknown protein shares substantial sequence homology with a known protein structure. Then, the structural pieces of the known protein can be used as a template to construct or piece together the unknown structural pieces of a target protein. For homology modeling, a large similarity between the template and target sequences is desired, and typically a sequence identity of >50% ensures that most of the structural features of a homology model will be accurate.

As Prox1 homeo-Prospero domain has 66.25% sequence identity with the *Drosophila* Prospero homeo-Prospero domain, we predicted the structure of Prox1 homeo-Prospero domain (580Leu-733Glu) using Prospero as template through homology modeling. There are in fact two available structures of Prospero, an apo form and a DNA bound form. Both structures were used to model Prox1 HDPD; however since the resulting structures were almost identical, we only report the structure modeled using the apo form of Prospero. The modeled structure was energy minimized (2000 steps steepest descent with cvff force field).

All of the secondary-structural elements, as well as the overall shape were preserved in the model structure of Prox1. Prox1, like Prospero, has a well-defined homeo-Prospero domain composed of three alpha-helices and an N-terminal extension. Helices I and II lie parallel to each other and across from Helix III. The third helix (part of the recognition helix) lies at the junction of the homeo-Prospero domain and is essential for DNA recognition and binding. Thus the homeo and Prospero domains act as one structural unit.



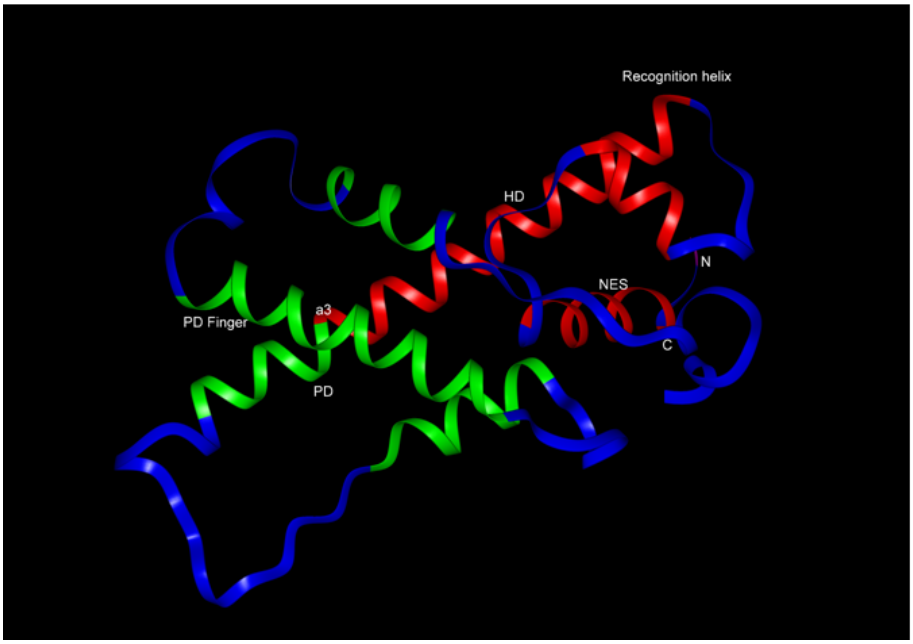


Figure 21: Predicted structure of Prox1 homeo-Prospero domain. Important structural features include three alpha helices characteristic of homeodomain proteins with two helices lying parallel and the third across from the other two (denoted as a3). Homeodomain (HD) is shown in red, and Prospero domain in green.

5.2 Validation of predicted Prox1 structure

Following energy minimization and refinement of the structure obtained by homology modeling, we assess the validity of this protein structure.

First, the geometric parameters (backbone torsional angles) of the model are viewed using the Ramachandran plot generated by RAMPAGE (S.C. Lovell 2002). More than 99% of the residues are classified in the allowed regions, leaving one out of 154 residues in a disallowed region. Furthermore, there are extensive alpha helices present with 102 residues belonging to the alpha helical regions and 18 residues in 3-10 helical regions.

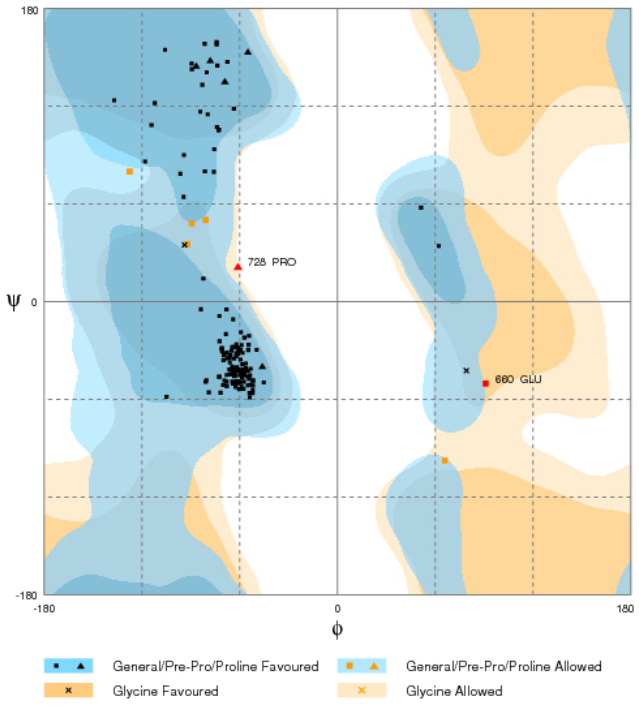


Figure 22: Ramachandran plot of Prox1 shows 99% of the residues are in the allowed regions

Next, the local model quality is assessed by plotting energies as a function of amino acid sequence position *i*. In general, positive values correspond to problematic parts of the input structure. Since a plot of single residue energies usually contains large fluctuations and is of limited value for model evaluation, the plot is smoothed by calculating the average energy over each 40-residue fragment. This calculation shows that almost all the residues are in energetically favorable local environment.

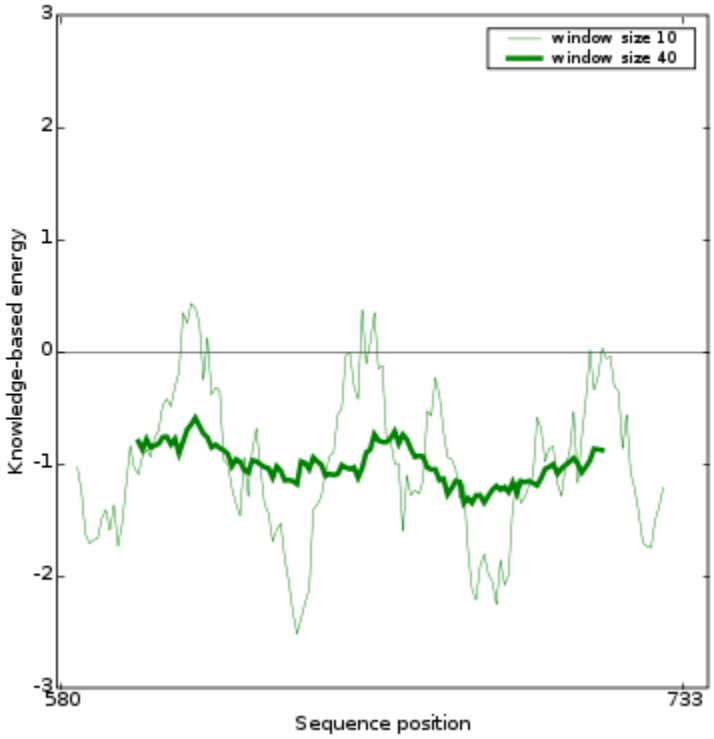


Figure 23: Energy as a function of amino acid position

Finally, we superimposed the predicted Prox1 structure on the Prospero template; these two structures are highly superimposable with an overall RMSD of only 1.1Å. These results indicate that the predicted model is a good workable structure for use in further analysis.

Chapter 6

STRUCTURE PREDICTION OF PCNA PROX1 COMPLEX

The yeast two-hybrid screen to detect Prox1 interacting partners produced 125 positive colonies corresponding to 26 unique proteins, including sc35 that was discussed above. Another interesting protein in the list was PCNA, proliferating cell nuclear antigen. It has been reported that the over-expression of Prox1 in blood endothelial cells induces upregulation of cell proliferation markers including PCNA, cyclin E1 and E2 (Petrova, Makinen et al. 2002). However in the lens, loss of Prox1 function leads to down-regulation of the cell cycle inhibitors $P27^{KIP1}$ and $P57^{KIP2}$ (Wigle, Chowdhury et al. 1999). Thus, it is plausible that Prox1 may function in both cell cycle arrest and cell cycle progression depending on interaction with these cell cycle regulators. The yeast two-hybrid screen did in fact identify PCNA as a Prox1 interacting protein. PCNA-Prox1 interaction has been studied extensively by Chen, in our lab, and here I only discuss the structural basis of this interaction that helps explain some of the experimental data collected by Chen.

6.1 Introduction to PCNA

PCNA is perhaps more famously known as the eukaryotic DNA sliding clamp whose main function during DNA replication is to slide along the DNA molecule and act as a processivity factor for DNA modifying enzymes, including DNA polymerase δ (Kelman 1995; Hingorani 2000). It was initially identified as an auto-antigen in patients with systemic lupus erythematosis (Miyachi, Fritzler et al. 1978), and later it was defined as an essential factor for DNA replication, DNA repair, cell cycle control and chromatin remodeling (Kelman 1997). The diverse functions of PCNA are dependent on its unique ring-shaped structure called "sliding clamp". PCNA forms a trimeric complex with six-fold symmetry in which the outside surface is composed of β-sheets and parallel α-helices and the inner surface of the hole is perpendicular to the DNA backbone. Thus PCNA can encircle double strand DNA and slide freely along it, so proteins can touch the DNA via interacting with the PCNA outer surface (Schurtenberger, Egelhaaf et al. 1998). A monomer of PCNA has three structural domains. The N-terminal and C-terminal domains are held together by a small interdomain connecting loop (IDCL) which serves as a binding site for most PCNA interacting proteins.

There are close to 30 proteins, involved in DNA repair, DNA replication, apoptosis, and cell cycle progression, that bind to PCNA. In most PCNA interacting proteins, the principal interaction is mediated by the recognition of a PCNA-interacting protein motif (PIP-box) of the interacting enzyme, which binds the inter-domain connector loop (IDCL) of a PCNA subunit (Gulbis 1996). The PIP-box has the

consensus sequence Q-x-x-(h)-x-x-(a)-(a), where h represents any one of L/I/M hydrophobic amino acids; a represents aromatic amino acids F or Y; and x is any residue.

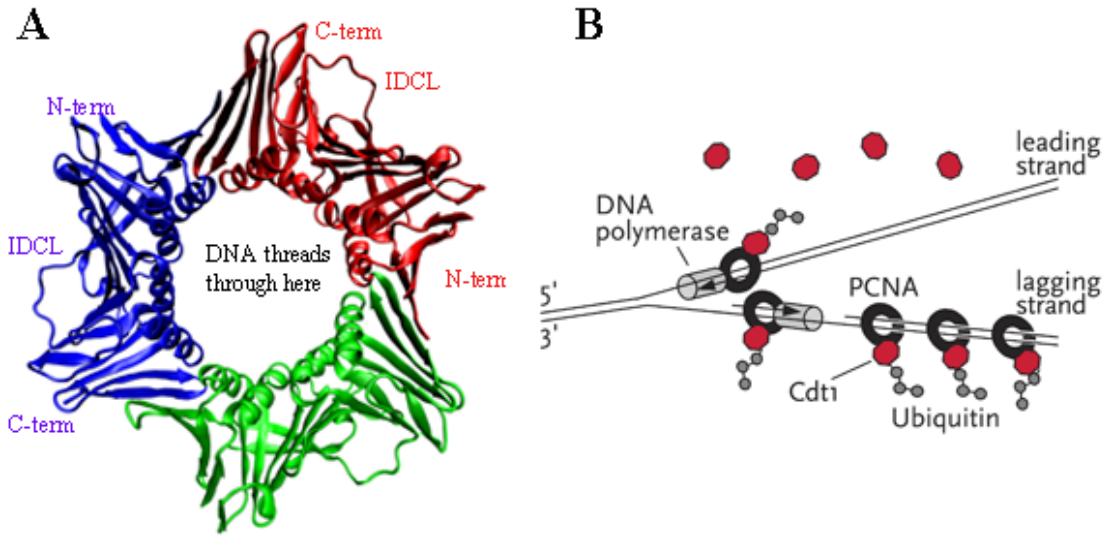


Figure 24: A: Structure of trimeric doughnut shaped PCNA molecule with N-terminal, C-terminal domains and the interdomain connecting loop (IDCL) labeled (pdb accession code 1AXC). B: Role of PCNA in DNA replication. The molecular doughnut encircles DNA like shower curtain rings on a rod, tethering the polymerase to the DNA. Besides polymerase, at least 30 other proteins involved in apoptosis, DNA repair, and cell cycle progression also bind to PCNA.

PCNA interaction with Prox1 has shown to repress Prox1 mediated activation of the chicken βB1 crystallin promoter (CAT assays done by Chen, data not shown). However, since multiple equivalent binding sites are present on the oligomeric toroidal clamp and a large number of possible PCNA clients that could be recruited to the limited binding sites on PCNA, this raises the question as to how assembly of functionally 'meaningful' combinations of clamp-bound enzymes is

regulated. Also, there will likely by competition among PCNA interacting proteins for access to PCNA and the winner determines which type of DNA modification will occur. Thus, even though Prox1 has been shown to interact with PCNA in yeast and CHO cells, it might be out-competed by other proteins in a living system. Furthermore, it is unclear how PCNA represses Prox1 function. One possibility is that PCNA interaction blocks the DNA binding domain of Prox1. Also puzzling is the fact that PCNA only forms the doughnut shaped trimer during DNA synthesis so we are unsure whether Prox1 directly interacts with the trimeric PCNA (in which case we have found a new function of Prox1 in DNA regulation rather than the previously reported role in gene regulation), or the monomeric PCNA subunit (in which case PCNA acts to regulate Prox1 function). In order to answer these questions, we propose a structural model for Prox1-PCNA complex by docking the solved PCNA structures (both trimeric and monomeric) on the predicted Prox1 HDPD structure.

6.2 Proposed structures for Prox1-PCNA complex

A sequence scan of Prox1 revealed that it too has a PIP-box motif in the Prospero domain, 686-**Q**-I-T-**L**-R-E-**F**-**F**-693 (the consensus residues are in boldface). Thus, it is very likely that Prox1 interacts with PCNA, like many other PCNA interacting proteins via the PIP-box. A notable feature of the PIP-box is that it folds into a tight 3_{10} helix with hydrophobic residues facing the surface. While the inter-domain connecting loop of PCNA presents a rather linear extension of non-polar

residues. These features make it possible for PCNA interacting proteins to bind PCNA by inserting the hydrophobic key into a non-polar pocket. Our proposed model for Prox1 structure also contains the PIP-box in a tight 3_{10} helix. Thus, we tested whether the interaction through the PIP-box is compatible with Prox1 and PCNA structures by docking Prox1 molecule onto PCNA. During docking, we also require that the PIP-box be contained in the interface region. Two residues, belonging to different proteins, are defined to be interacting if the distance between at least two atoms (of the two residues) is smaller than 6Å. In this manner, we propose two models, one in which full length trimeric PCNA is used to dock with Prox1, and in the other, only the C-terminus domain of PCNA (167-255) is used. The reason for using the C-terminal PCNA structure is because the positive hit for PCNA from the original yeast two-hybrid contained a DNA fragment that corresponds to the C-terminal PCNA domain.

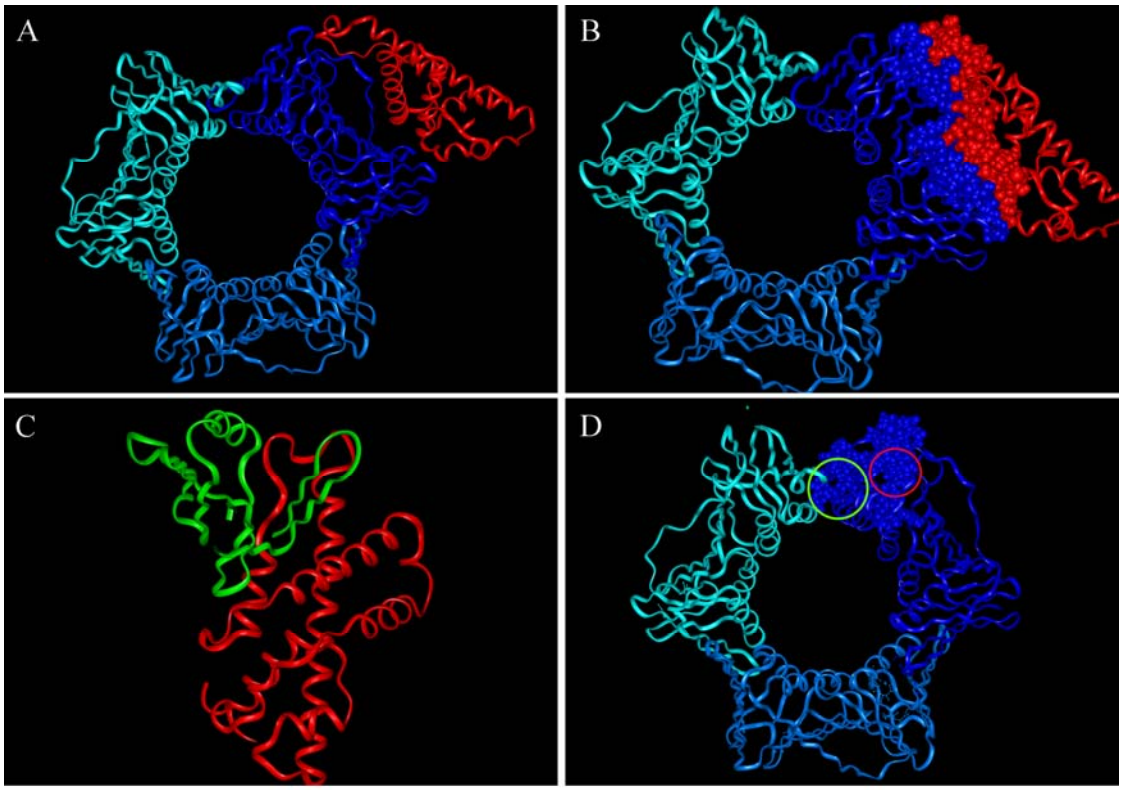


Figure 25: Two structures for Prox1-PCNA interaction. A: Ribbon diagram of the full length trimeric PCNA-Prox1 complex. Prox1 is colored red in all subsequent figures; B: Interface residues from A are depicted as van der Walls spheres; C: Ribbon diagram of the interaction between the C-terminus of PCNA (green) and Prox1; D: PCNA molecule showing the interface residues from C as van der Walls spheres.

The two proposed models are depicted in figure above; the importance of panel D is discussed later. The buried surface area upon complex formation was also computed for both models. In trimeric PCNA-Prox1 complex, 2150Å^2 surface area is buried whereas 1720Å^2 is buried for C-term PCNA/Prox1 complex. Although the absolute buried surface area is higher with trimeric PCNA/Prox1, it only corresponds to roughly 20% of the entire complex surface area; whereas, in the case of C-terminal PCNA/Prox1 complex, nearly 28% of the entire surface area is buried. Thus, there are relatively more residues making contacts with Prox1 in the latter case.

6.3 Discussion

Through docking simulations, we have shown that Prox1 can in fact bind to PCNA through the PIP-box – the geometry of both proteins is fitting to allow substantial interactions if the PIP-box is forced to lie in the interface region. This theory has been experimentally validated by site-directed mutagenesis of the PIP-box one residue at a time (experiment performed by Chen; data not shown). Chen found that of the PIP-box (686-Q-I-T-L-R-E-F-F-693), the first phenylalanine residue (Phe692) was crucial for interaction, whereas mutation of the second phenylalanine (Phe693) resulted in little change. Furthermore, the residues glutamine and leucine were also deemed important, although not as significant as Phe692. This result was actually predicted from our model of PCNA bound Prox1 structure (panel A of fig 23). Although the PIP-box forms a tight 3_{10} helix, the residue Phe693 actually faces the inner surface of the protein and is inaccessible to PCNA for interaction. However, Phe692, Glu686, and Leu689 are positioned for direct access to PCNA.

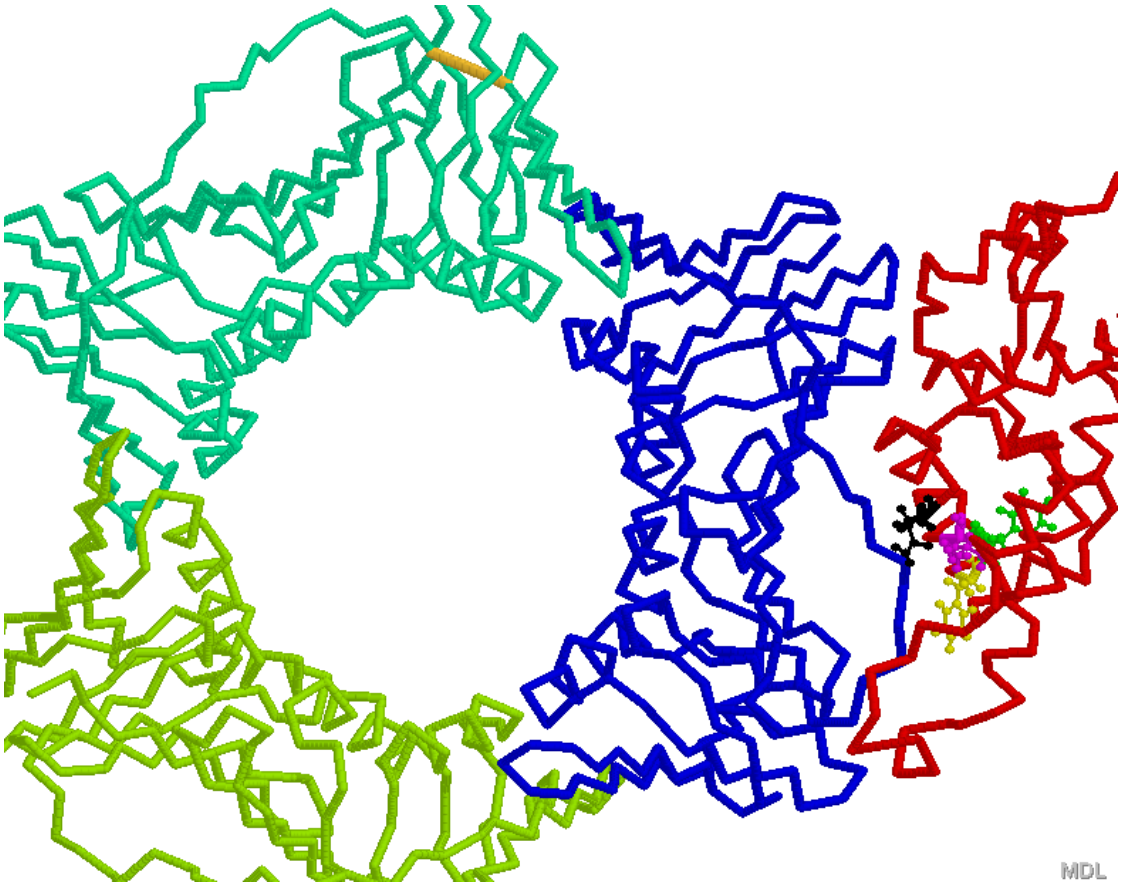


Figure 26: PCNA/Prox1 complex represented as a backbone trace; Prox1 is in red. The residue Phe693, shown in green, faces away from the surface and is inaccessible for interaction with PCNA; whereas residues Phe692 (yellow), Glu686 (black) and Leu689 (magenta) are well positioned for direct interaction with the IDCL of PCNA.

Furthermore, our model predicts that Prox1 can bind to either the C-terminus or the IDCL of PCNA. The favorable interaction afforded by C-terminus alone is, however, not entirely realistic for two reasons. First, since PCNA trimerizes by associating C-terminus of one monomer to the N-terminus of another, most of the C-terminal residues will become inaccessible to Prox1 when PCNA trimerizes. These inaccessible residues are highlighted with a green circle in panel D of fig 24. Secondly, if we extend the C-terminus to include the inter-domain connecting loop,

we find that some of the previously interacting residues have now become inaccessible for binding to Prox1 because the IDCL masks their positions. These residues are highlighted with a red circle in panel D of fig 24. However, the upshot is that the contacts lost by the masking of the C-terminus are compensated for by IDCL itself providing a new surface to contact with Prox1. Thus, in the end, Prox1 interacts mainly with the inter-domain connecting loop with some extension into the C-terminus for trimeric PCNA structure. In the monomer, the C-terminal residues involved in trimerization are now accessible for binding so there is a stronger interaction; however, the interface region is still confined to IDCL and the C-terminus. This model is consistent with the fact that there was not a huge difference in absolute buried surface area between the trimeric PCNA/Prox1 and C-term PCNA/Prox1 complexes since contacts lost by IDCL masking are compensated for in the trimeric structure. Furthermore, yeast two-hybrid using different truncations of PCNA (N-terminus alone, N-terminus with IDCL, C-terminus alone, C-terminus with IDCL and full length PCNA) have confirmed that interaction with Prox1 does not occur with the N-terminus of PCNA and is possible only with C-terminus or IDCL (experiments performed by Chen; data not presented).

Inspired by docking simulations and validated by yeast two-hybrid, we have now determined that Prox1 binds to the same region of PCNA as the other 30 PCNA-interacting proteins. And although Prox1-PCNA interaction is not very strong, it might be sufficient to out-compete some other proteins for access to PCNA, but in

order to test this hypothesis, one would need to compute the binding free energy of interaction either computationally through a method such as MM-PBSA or experimentally. This will be done in the future. However, even if it turns out that Prox1 cannot out-compete other PCNA interacting proteins, it may still be able to bind the monomer PCNA subunit. Most protein interactions with PCNA occur when it is trimerized and has encircled a DNA molecule. Since our simulations predict that Prox1 can bind the monomeric form more readily, Prox1-monomer PCNA might be the primary mode of interaction within the cell nucleus.

The significance of this interaction is very interesting. Note that cellular differentiation is induced by Prox1, at least in the lens; and cell proliferation cannot occur without high levels of PCNA. And since cell differentiation is usually preceded by cell proliferation, there may be a strong link between PCNA and Prox1 in coordinating the timing of differentiation. Also, the decision to proliferate or differentiate may be influenced by the relative amounts of Prox1 and PCNA in the cell. The high levels of PCNA in proliferating cells inhibit Prox1; when proliferation ceases and PCNA levels decrease, Prox1 is no longer repressed and begins its program of cell fate determination.

In any event, the DNA binding site on Prox1 (part of the homeodomain) is completely removed from the regions that bind PCNA (part of Prospero domain). So, the observed repression of Prox1 activity upon PCNA interaction is not due to physical blockage of the DNA recognition helix. The explanation may be simpler such

as: Prox1 becomes inefficient at binding DNA when it is carrying a huge load along with it; or that PCNA binding induces a slight conformational change that interferes with DNA recognition.

Chapter 7

STRUCTURAL BASIS FOR PROX1 DNA INTERACTION

7.1 Prox1 interaction with the chicken βB1-crystallin promoter

Prox1 is known to bind to the OL2 element of the chicken βB-1 crystallin promoter and activate its gene transcription. Our lab has recently found two new Prox1 binding sites in the chicken βB1 crystallin promoter, called the -220 and -290 sites. DNAase I foot-printing and gel shift assays have shown that Prox1 binds strongly to the -220 site and represses the chicken βB1 crystallin promoter. Repression is also observed with the -290 site, although not as severe as the -220 site (data not shown).

7.2 Structural models of Prox1 bound DNA

We used the previously reported OL2 site, and the newly discovered -220 and -290 sites to determine whether there is a good geometric and chemical compatibility with the Prox1 recognition helix. A good fit both validates the predicted structure of Prox1 and further confirms the presence of two new Prox1 binding sites in the chicken βB1 crystallin promoter. The structures of Prox1/DNA complex were generated by docking a short fragment of DNA onto Prox1 structure.

Our docking model shows that both OL2 and -220 sites bind to Prox1 through a critical AG dinucleotide. This AG dinucleotide resides in the major groove of DNA and three conserved residues Lys622, Asn626, and Glu629 make sequence specific hydrogen bond contacts. Flanking nucleotides make van der Walls or hydrophobic contacts with the recognition helix and further stabilize the complex. It is interesting to note that the three residues responsible for making hydrogen bond contacts with AG dinucleotide are also conserved in Prospero where they carry out similar function in TDA pros DNA binding. The -290 site however does not bind through the AG nucleotide, rather a cytosine nucleotide directly upstream AG is predicted to be important for recognition.

In comparing the three structures, there are noticeably more contacts between Prox1 and the -220 site than the other two sites. The -290 site is very similar in sequence to the OL2 site, however, -290 makes more contacts with Prox1 than OL2, but less than -220. Furthermore, the buried surface area of the interface region was calculated to be 1800Å^2 (20% of the total surface area) for Prox1-220 complex; 1705Å^2 (17.5% of the total surface area) for Prox1-290 complex; and 1620Å^2 (14% of the total surface are) for Prox1-OL2 complex. Taken together, these data suggest that Prox1 binds most tightly at the -220 site, then -290 and least tightly to the OL2 site. Of course we realize that such a conclusion is not always warranted and that the number of contacts and buried surface area cannot alone dictate binding affinity; however, the

present data are highly suggestive of binding order -220>-290>OL2. This predicted trend correlates well with experimental data.

The most likely orientation of Prox1 bound to the OL2 site is depicted in fig 26; similar poses were obtained for Prox1-220/-290 sites (data not shown). Note that we have used canonical B-form DNA structures to model the interaction and have assumed both ligand and receptor to be rigid bodies during docking. Prox1 can be assumed to be rigid by analogy to Prospero. The experimentally determined native and DNA bound Prospero structures are nearly identical in the recognition helix region and the two structures have an overall RMSD of only .94Å due to a slight shift near the NES region upon binding DNA. Thus, we assume that Prox1 will too behave similarly when bound to DNA and will not exhibit much change in the recognition helix. However, the local structure of the bound DNA may change upon protein binding. This does not pose a big problem for us because we are only interested in comparing the orientations of DNA and Prox1 and inferring the relative binding affinities from the number of contacts. Since cranking the DNA structure is energetically unfavorable, the structure would likely change only if the energy stabilization afforded by new protein-DNA contacts outweigh the expenditure to change the DNA structure. Thus, even if the DNA structure changes during binding, the relative order of protein-DNA contacts is still maintained.

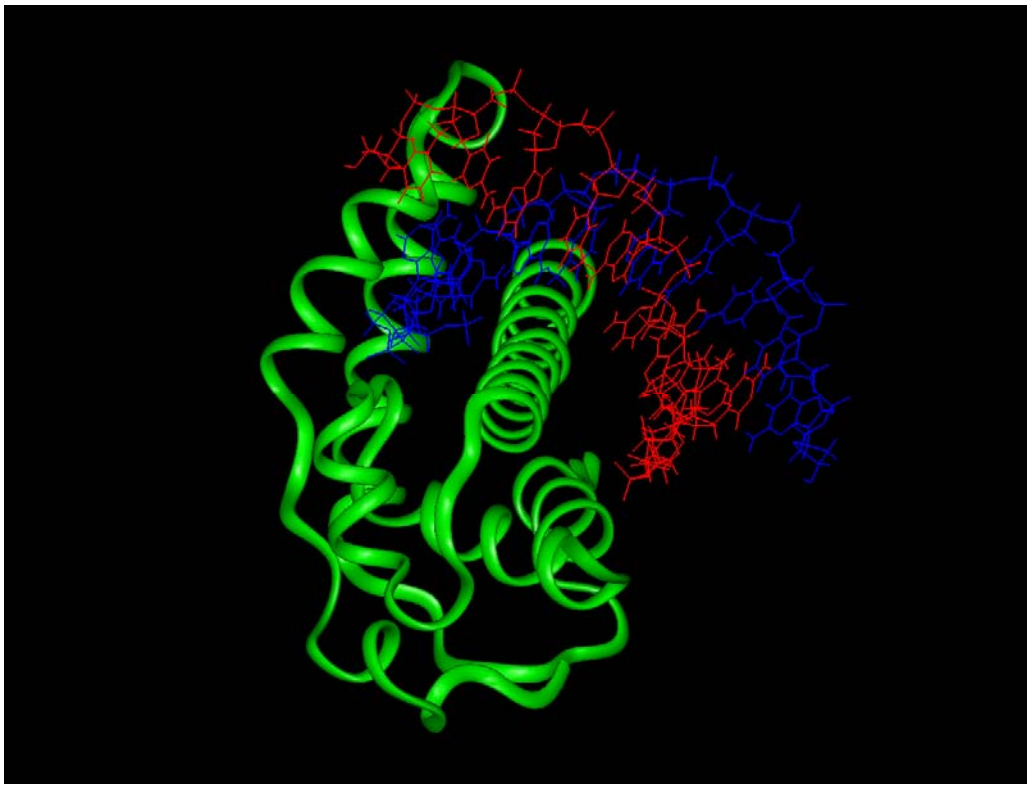


Figure 27: Model of Prox1 bound to the OL2 site. The Prox1 HDPD is shown in green; DNA makes contacts with the helix III of homeodomain.

Thus, these structural models provide further insight into Prox1-DNA interaction.

Chapter 8

CONCLUSION

Prox1 is a multifunctional transcription factor that plays many different roles in tissue specific manner. The function of such an important protein must be tightly controlled, especially during development. We already have evidence that loss of functional control over Prox1 leads to cancerous cell types. In a traditional sense, protein control can be achieved on the level of gene transcription, posttranslational modifications or by specific protein-protein interactions. The former two cases are likely to play minor roles in the functional regulation of Prox1, since (i) tissues that normally contain Prox1 have relatively similar amounts of Prox1 protein expression and (ii) glycosylation, or phosphorylation do not occur in Prospero and such modifications have not yet been reported for Prox1 either. In addition, Prox1 has been reported to interact with nuclear receptors involved in maintaining lipid homeostasis and loss of Prox1 function leads to loss of cell cycle control. These data then highly suggest that Prox1 function is modulated by protein-protein interactions more than any other route.

Indeed, our lab found 26 unique nuclear proteins that can interact with Prox1 via a library-scaled yeast two-hybrid. One protein in particular, sc35 interacted with Prox1 multiple times in our screen, suggesting a prominent interaction. Further

experiments determined that the splicing factor sc35 represses Prox1's function as a transcription factor in a concentration dependent manner. However, Prox1-sc35 interaction is not as simple as it may seem. Since sc35 is usually localized to the storage facilities, the nuclear speckles, and Prox1 is no where to be found within the speckles, this raises the question as to how do these proteins interact if they are compartmentalized in different regions of the cell nucleus. Through a series of CAT assays and immunofluorescence data, we now believe that Prox1-sc35 interaction is only present when the cell is experiencing high transcriptional activity. In such an event, splicing factors, including sc35, diffuse out of the nuclear speckles and are recruited to the site of pre-mRNA synthesis. Only during this moment, Prox1 and sc35 are present in the same sub-compartment and are able to interact. Thus, in a way, sc35 acts like a molecular brake by repressing Prox1 mediated gene activation during high transcription periods. This is perhaps one way the cell keeps Prox1 (and gene transcription) in check.

Similarly, Prox1 function is also modulated by its interaction with the proliferating cell nuclear antigen, PCNA. The structural models of Prox1-PCNA complex have illuminated how these two proteins interact and partly explain how PCNA represses Prox1 function. This specific interaction may be important to control the timing of differentiation or may play a role in the decision to proliferate or differentiate during early stages of development.

These two proteins, sc35 and PCNA, are just a subset of many Prox1 interacting proteins, each one of which has the potential to change the function of Prox1. Thus, in addition to Prox1 function regulation by competition of other transcription factors such as PAX6 and E2F to the same DNA binding site, we report that Prox1 function can also be modulated by protein-protein interactions. The latter may be an important mode for Prox1 regulation especially since protein-protein interactions are highly specific and cell-type dependent allowing finer spatio-temporal control.

References:

A. D. MacKerell, J., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus (1998). "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins " J Phys Chem **102**(18): 3586-3616.

Al-Rawi, M. A., R. E. Mansel, et al. (2005). "Lymphangiogenesis and its role in cancer." Histol Histopathol **20**(1): 283-98.

Banerjee-Basu, S., D. Landsman, and A.D. Baxevanis (1999). "Threading analysis of prospero-type homeodomains." In silico Biol. **1**(3): 163-73.

Berger, C., Pallavi, S.K., Prasad, M., Shashidhara, L.S., Technau, G.M. (2005). "A critical role for cyclin E in cell fate determination in the central nervous system of Drosophila melanogaster." Nat Cell Biol **7**(1): 56-62.

Brenton R Graveley, a. K. J. H. (2005). SR Proteins. Encyclopedia of life sciences, John Wiley & Sons.

Burglin, T. R. (1994). "A Caenorhabditis elegans prospero homologue defines a novel domain." Trends Biochem Sci **19**(2): 70-1.

Burke, Z. a. G. O. (2002). "Prox1 is an early specific marker for the developing liver and pancreas in the mammalian foregut endoderm." Mech Dev **118**(1-2): 147-55.

Carmeliet, K. A. a. P. (2002). "Molecular mechanisms of lymphangiogenesis in health and disease." Cancer Cell **1**: 219-227.

Chen R, L. L., Weng Z (2003). "ZDOCK: An Initial-stage Protein-Docking Algorithm." Proteins **52**: 80-87.

Chen R, W. Z. (2002). "Docking Unbound Proteins Using Shape Complementarity, Desolvation, and Electrostatics." Proteins **47**: 281-294.

Chen R, W. Z. (2003). "A Novel Shape Complementarirty Scoring Function for Protein-Protein Docking." Proteins **51**: 397-408.

Chen X, P. T., Cain W, and Duncan MK (2006). "Production of monoclonal antibodies against Prox1." Hybridoma (Larchmt) **25**(1): 27-33.

Craig, D. P. a. J. A. (1974). "Cataracts: Development, diagnosis, and management " CIBA Clin. Symp. **26**(3): 2-32.

Cui, W., S. I. Tomarev, et al. (2004). "Mafs, Prox1, and Pax6 can regulate chicken betaB1-crystallin gene expression." J Biol Chem **279**(12): 11088-95.

Czerny T, Halder G, et al. (1999). "twin of eyeless, a second Pax-6 gene of Drosophila, acts upstream of eyeless in the control of eye development." Mol Cell. **3**(3): 297-307.

Doe, C. Q. (1992). "Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system." Development **116**: 855-863.

Doe, C. Q., Q. Chu-LaGraff, et al. (1991). "The prospero gene specifies cell fates in the Drosophila central nervous system." Cell **65**(3): 451-464.

Dudas J, Papoutsi M, et al. (2004). "The homeobox transcription factor Prox1 is highly conserved in embryonic hepatoblasts and in adult and transformed hepatocytes, but is absent from bile duct epithelium." Anat Embryol (Berl) **208**(5): 359-66.

Duncan MK, Cui W, et al. (2002). "Prox1 is differentially localized during lens development." Mech Dev **112**(1-2): 195-8.

Duncan, M. K., W. Cui, et al. (2002). "Prox1 is differentially localized during lens development." Mech. Dev. **112**: 195-198.

Duronio, D. L. M. a. R. J. (2000). "Cell cycle: To differentiate or not to differentiate?" Curr. Biol. **10**(8): R302-R304.

Farhad Abtahian, A. G., Eric Sebzda, Min-Min Lu, Rong Zhou, Attila Mocsai, Erin E. Myers, Bin Huang, David G. Jackson, Victor A. Ferrari, Victor Tybulewicz, Clifford A. Lowell, John J. Lepore, Gary A. Koretzky, Mark L. Kahn2dagger (2003). "Regulation of Blood and Lymphatic Vascular Separation by Signaling Proteins SLP-76 and Syk." Science **299**(5604): 1079477.

Ge, H. a. J. L. M. (1990). "A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro." Cell **62**: 25-34.

Gehring, W. J., M. Affolter, and T. Burglin (1994). "Homeodomain proteins." Annu Rev Biochem **63**: 487-526.

Grainger, R. M. (1992). "Embryonic lens induction: shedding light on vertebrate tissue determination." Trends Genet **8**(10): 349-55.

Gruss, R. A.-P. a. P. (2001). "Pax6 lights-up the way for eye development." Curr Opin Cell Biol **13**(6): 706-14.

Guex, N. a. P., M. C. (1997). "SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling." Electrophoresis **18**: 2714-2723.

Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., Kuriyan, J. (1996). "Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA." Cell **87**: 297-306.

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

Hingorani, M. M. a. O. D., M. (2000). "Sliding clamps: a (tail)ored fit." Curr. Biol. **10**: R25-R29.

Hirata, J., Nakagoshi, H., Nabeshima, Y., Matsuzaki, F. (1995). "Asymmetric segregation of the homeodomain protein Prospero during *Drosophila* development." Nature **377**: 627-630.

Hogan, Y. F. a. B. L. M. (1998). "BMP4 is essential for lens induction in the mouse embryo." Genes Dev **12**: 3764-3775.

Hong YK, Harvey N, et al. (2002). "Prox1 is a master control gene in the program specifying lymphatic endothelial cell fate." Dev Dyn **225**(3): 351-7.

Huang, S. a. D. L. S. (1996). "Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription. ." J. Cell Biol. **133**: 719-732.

Hubbard, S. J. T., J. M. (1993). "NACCESS: atomic solvent accessible area calculations."

Jan, Y. N., Jan, L.Y (1998). "Asymmetric cell division." Nature **392**: 775-778.

Jiménez-García, L. F. a. D. L. S. (1993). "In vivo evidence that transcription and splicing are coordinated by a recruiting mechanism." Cell **73**: 47-59.

Jodi M. Ryter, C. Q. D. a. B. W. M. (2002). "Structure of the DNA Binding Region of Prospero Reveals a Novel Homeo-Prospero Domain." Structure **10**(11): 1541-1549.

Jun Qin, D.-m. G., Quan-Feng Jiang, Qing Zhou, Yu-Ying Kong, Yuan Wang and You-Hua Xie (2004). "Prospero-Related Homeobox (Prox1) Is a Corepressor of Human Liver Receptor Homolog-1 and Suppresses the Transcription of the Cholesterol 7-{alpha}-Hydroxylase Gene." Mol Endocrinol **18**(10): 2424-39.

Kao, F. T. a. T. T. P. (1968). "Genetics of somatic mammalian cells, VII. Induction and isolation of nutritional mutants in Chinese hamster cells." Proc Natl Acad Sci USA **60**(4): 1275-81.

Keith L. Parker, D. A. R., Deepak S. Lala, Yayoi Ikeda, Xunrong Luo, Margaret Wong, Marit Bakke, Liping Zhao, Claudia Frigeri, Neil A. Hanley, Nancy Stallings and Bernard P. Schimmer (2002). "Steroidogenic Factor 1: an Essential Mediator of Endocrine Development." Recent Prog Horm Res **57**: 19-36.

Kelman, Z. (1997). "PCNA: structure, functions and interactions." Oncogene **14**(6): 629-40.

Kelman, Z. a. O. D., M. (1995). "DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine." Annu. Rev. Biochem. **64**: 171-200.

Knoblich, J. A., Jan, L.Y., Jan, Y.N. (1995). "Asymmetric segregation of Numb and Prospero during cell division." Nature **377**: 624-627.

Krainer, A. R., G.C. Conway, and D. Kozak (1990). "The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites." Cell **62**: 35-42.

Krainer, A. R., G.C. Conway, and D. Kozak (1990). "Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells." Genes & Dev **4**: 1158-1171.

Laerm A, Helmbold P, et al. (2007). "Prospero-related homeobox 1 (PROX1) is frequently inactivated by genomic deletions and epigenetic silencing in carcinomas of the bilary system." J Hepatol **46**(1): 89-97.

Lengler J, K. E., Tomarev S, Prescott A, Quinlan RA, Graw J (2001). "Antagonistic action of Six3 and Prox1 at the gamma-crystallin promoter." Nucleic Acids Research **29**(2): 515-26.

Li L, C. R. j. f. a., Weng Z (2003). "RDOCK: Refinement of Rigid-body Protein Docking Predictions." Proteins **53**: 693-707.

M.T. Mas, K. C. S., D.L. Yarmush, K. Aisaka and R.M. Fine (1992). "Modeling the anti-CEA antibody combining site by homology and conformational search,." Protein Struct Funct Genet **14**: 483-498.

McAvoy, J. W. a. C. G. C. (1989). "Fibroblast growth factor (FGF) induces different responses in lens epithelial cells depending on its concentration." Development **107**(2): 221-228.

Mintz, P. J., Patterson, S. D., Neuwald, A. F., Spahr, C. S., and Spector, D. L. (1999). "Purification and biochemical characterization of interchromatin granule clusters." Embo J **19**: 4308-20.

Misteli, T., Caceres, J. F., and Spector, D. L. (1997). "The dynamics of a pre-mRNA splicing factor in living cells." Nature **387**: 523-527.

Miyachi, K., M. J. Fritzler, et al. (1978). "Autoantibody to a nuclear antigen in proliferating cells." J Immunol **121**(6): 2228-34.

Murtaugh, L. C. a. D. A. M. (2003). "Genes, signals, and lineages in pancreas development." Annu Rev Cell Dev Biol **19**: 71-89.

Myster, D. L., Duronio, R.J. (2000). "To differentiate or not to differentiate?" Curr Biol **10**: R302-304.

Nagai H, Li Y, et al. (2003). "Mutations and aberrant DNA methylation of the PROX1 gene in hematologic malignancies." Genes Chromosomes Cancer **38**(1): 13-21.

Oliver G and D. M. (2002). "The rediscovery of the lymphatic system: old and new insights into the development and biological function of the lymphatic vasculature." Genes Dev **16**(7): 773-83.

Padera, R. K. J. a. T. P. (2003). "Lymphatics make the break." Science **299**: 209-210.

Petrova, T. V., T. Makinen, et al. (2002). "Lymphatic endothelial reprogramming of vascular endothelial cells by the Prox-1 homeobox transcription factor." Embo J **21**(17): 4593-9.

Piatigorsky, J. (1981). "Lens differentiation in vertebrates. A review of cellular and molecular features." Differentiation **19**(3): 134-53.

Pierre Val, A.-M. L.-M., Georges Veyssière, and Antoine Martinez (2003). "SF-1 a key player in the development and differentiation of steroidogenic tissues." Nucl Recept **1**(1): 8.

R. Ashery-Padan, T. M., X. Zhou and P. Gruss (2000). "Pax6 activity in the lens primordium is required for lens formation and for correct placement of a single retina in the eye." Genes Dev **14**: 2701–2711.

R.A. Laskowski, M. W. M., D.S. Moss and J.M. Thornton (1993). "PROCHECK: a program to check the stereochemistry of protein structures." J Appl Cryst **26**: 283-291.

Rajarshi Maiti, G. H. V. D., Haiyan Zhang, and David S. Wishart (2004). "SuperPose: a simple server for sophisticated structural superposition." Nucleic Acids Research **32**(Web server issue W590W594).

Reed NA, O. D., Czymmek KJ, Duncan MK (2001). "An immunohistochemical method for the detection of proteins in the vertebrate lens." J Immunol Methods **253**: 243-252.

Reis-Filho, J. S. and F. C. Schmitt (2003). "Lymphangiogenesis in tumors: what do we know?" Microsc Res Tech **60**(2): 171-80.

Richards, J. W. P. a. F. M. (1987). "Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes,." J Mol Biol **193**: 775-791.

S.C. Lovell, I. W. D., W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson and D.C. Richardson (2002). "Structure validation by Calpha geometry: phi,psi and Cbeta deviation." Proteins: Structure, Function & Genetics **50**: 437-450.

Saha MS, S. C., Grainger RM. (1989). "Embryonic lens induction: more than meets the optic vesicle." Cell Differ Dev **28**(3): 153-71.

Schurtenberger, P., S. U. Egelhaaf, et al. (1998). "The solution structure of functionally active human proliferating cell nuclear antigen determined by small-angle neutron scattering." J Mol Biol **275**(1): 123-32.

Schwede, J. K. a. T. (2004). "The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models " Nucleic Acids Research **32**: D230-D234.

Schwede T, K. J., Guex N, and Peitsch MC (2003). "SWISS-MODEL: an automated protein homology-modeling server." Nucleic Acids Research **31**: 3381-3385.

Slack, J. M. (1995). "Developmental biology of the pancreas." Development **121**(6): 1569-80.

Song KH, Li T, et al. (2006). "A Prospero-related homeodomain protein is a novel co-regulator of hepatocyte nuclear factor 4alpha that regulates the cholesterol 7alpha-hydroxylase gene." J Biol Chem **281**(15): 10081-8.

Sosa-Pineda, B., J.T. Wigle, and G. Oliver (2000). "Hepatocyte migration during liver development requires Prox1." Nat Genet **25**(3): 254-5.

Spana, E. P., Doe, C.Q. (1995). "The prospero transcription factor is asymmetrically localized to the cell cortex during neuroblast mitosis in *Drosophila*." Development **121**: 3187-3195.

Spector, A. I. L. D. L. (2003). "Nuclear speckles: a model for nuclear organelles." Nature reviews molecular cell biology **4**: 605-612.

Spector, D. L. (1993). "Macromolecular domains within the cell nucleus." Annu. Rev. Cell Biol. **9**: 265-315.

Steffensen KR, Holter E, et al. (2004). "Functional conservation of interactions between a homeodomain cofactor and a mammalian FTZ-F1 homologue." EMBO Rep **5**(6): 613-9.

Steffensen KR, Holter E, et al. (2004). "Functional conservation of interactions between a homeodomain cofactor and a mammalian FTZ-F1 homologue." EMBO Rep **5**(6): 613-9.

Tatiana V. Petrova, T. M., Tomi P. Mäkelä, Janna Saarela, Ismo Virtanen, Robert E. Ferrell, David N. Finegold, Dontscho Kerjaschki, Seppo Ylä-Herttuala and Kari Alitalo (2002). "Lymphatic endothelial reprogramming of vascular endothelial cells by the Prox-1 homeobox transcription factor." The EMBO Journal **21**(17): 4593-4599.

Thiry, M. (1995). "The interchromatin granules." Histol Histopathol **10**: 1035-45.

Tomarev, S. I., R. D. Zinovieva, et al. (1998). "Characterization of the mouse Prox1 gene." Biochem Biophys Res Commun **248**(3): 684-9.

Turner M, M. P., Costello PS, Williams O, Price AA, Duddy LP, Furlong MT, Geahlen RL, Tybulewicz VL. (1995). "Perinatal lethality and blocked B-cell development in mice lacking the tyrosine kinase Syk." Nature **378**(6554): 298-302.

Van der Auwera, I., S. J. Van Laere, et al. (2004). "Increased angiogenesis and lymphangiogenesis in inflammatory versus noninflammatory breast cancer by real-time reverse transcriptase-PCR gene expression quantification." Clin Cancer Res **10**(23): 7965-71.

Wang J, Kilic G, et al. (2005). "Prox1 activity controls pancreas morphogenesis and participates in the production of "secondary transition" pancreatic endocrine cells." Dev Biol **286**(1): 182-94.

Wenwu Cui, S. I. T., Joram Piatigorsky, Ana B. Chepelinsky, and Melinda K. Duncan (2004). "Mafs, Prox1, and Pax6 Can Regulate Chicken betaB1-Crystallin Gene Expression." THE JOURNAL OF BIOLOGICAL CHEMISTRY **279**(12): 11088-11095.

Wigle JT, Chowdhury K, et al. (1999). "Prox1 function is crucial for mouse lens-fibre elongation." Nat Genet **21**(3): 318-22.

Wigle JT and O. G (1999). "Prox1 function is required for the development of the murine lymphatic system." Cell **98**(6): 769-78.

Wigle, J. T., K. Chowdhury, et al. (1999). "Prox1 function is crucial for mouse lens-fibre elongation." Nat Genet **21**(3): 318-22.

Wilting, J., M. Papoutsi, et al. (2002). "The transcription factor Prox1 is a marker for lymphatic endothelial cells in normal and diseased human tissues." Faseb J **16**(10): 1271-3.

Yamada, T., T. Nakamura, et al. (1990). "Synthesis of alpha-crystallin by a cell line derived from the lens of a transgenic animal." Curr Eye Res **9**(1): 31-7.

Yi-Wen Liu, W. G., Hui-Ling Teh, Jee-Hian Tan, and Woon-Khiong Chan (2003). "Prox1 Is a Novel Coregulator of Ff1b and Is Involved in the Embryonic Development of the Zebra Fish Interrenal Primordium." Mol Cell Biol **23**(20): 7243-7255.

Yousef, M. S., Matthews, B.W. (2005). "Structural basis of Prospero-DNA interaction: implications for transcription regulation in developing cells." Structure **13**: 601-7.

Zahler, A. M., K.M. Neugebauer, W.S. Lane, and M.B. Roth (1993). "Distinct functions of SR proteins in alternative pre-mRNA splicing." Science **260**: 219-222.

Zoe¨ Burke, G. O. (2002). "Prox1 is an early specific marker for the developing liver and pancreas in the mammalian foregut endoderm." Mech Dev **118**: 147-155.