Two Splice Variants of Nopp140 in Drosophila Melanogaster.

John Mark Waggener
Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/369
INFORMATION TO USERS

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

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

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

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

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

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

UMI ®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
TWO SPLICE VARIANTS OF NOPP140 IN DROSOPHILA MELANOGASTER

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biological Sciences

by
John Mark Waggener
B.S., Louisiana State University, 1994
August 2001
Dedication

This dissertation is dedicated to my wife, Natalie. Natalie has always supported my decisions and been there when things were not going so well. Without her support and understanding I could never have reached this point.
Acknowledgments

I am indebted to many people for helping me through graduate school. I would like to first thank my graduate advisory committee, Drs. Simon Chang, Richard Bruch, and Grover Waldrop for fully supporting my research efforts and always treating me with respect. My graduate advisor, Dr. Patrick DiMario, has taught me how to always ask the next question, and I feel that he has trained me well. I am grateful for a piece of his knowledge of many diverse techniques.

Many others have helped me along the way by sharing what they had. The cDNA library made from *Drosophila* stage 10 egg chambers was a gift from Drs. John Tower and Allan Spradling (Department of Embryology, Carnegie Institution of Washington, Baltimore, MD). The Schneider II cell line was a gift from Dr. Allan Spradling. I thank Drs. Joseph Gall and Zheng'an Wu (Department of Embryology, Carnegie Institution of Washington, Baltimore, MD) for the cDNA encoding *Xenopus laevis* coilin. I thank Dr. Benigno (Ben) Valdez (Baylor College of Medicine, Houston, TX) for providing me with antibodies directed against human nucleolin and Dr. Robert Ochs (Scripts Research Institute, LaJolla, CA) for the anti-human S4 anti-fibrillarin antibody. I would also like to thank everyone on the fifth floor of Choppin Hall from whom I have ever begged, borrowed, or stolen anything. I also thank my lab mate Greg Pellar for showing me the ropes.

I would especially like to thank my parents. They instilled in me a love of reading and learning new things. I thank my brother Joel for his help with my figures. Without his design expertise, I do not know if I would have ever finished this document.
# Table of Contents

Dedication .................................................................................................................... iii

Acknowledgments ......................................................................................................... iv

List of Figures ............................................................................................................... vi

Abstract ....................................................................................................................... vii

Chapter 1  
Literature Review ...................................................................................................... 1

Chapter 2  
Identification of Two Splice Variants of Nopp140 in *Drosophila melanogaster*: Screening, Sequencing, and Comparison of Nopp140 Homologues Across Species ........................................ 29

Chapter 3  
Expression and Phosphorylation of DmNopp140 .............................................46

Chapter 4  
Localization Patterns of DmNopp140-RGG and DmNopp140 Within Nucleoli and Cajal bodies .......................................................... 57

Chapter 5  
Conclusions and Future Directions ................................................................. 93

Literature Cited .............................................................................................................102

Appendix A: Primer Locations................................................................................... 119

Appendix B: Vector Maps............................................................................................120

Vita............................................................................................................................... 130
### List of Figures

1.1 Maturation of the eukaryotic ribosome .......................................................... 3
1.2 The box H/ACA and box C/D snoRNAs ......................................................... 6
1.3 Overview of RNA pathways within the nucleolus .......................................... 9
1.4 Interactions of Noppl40 with box H/ACA and box C/D snoRNP....................... 27

2.1 Expressed sequence tags exist for both *D. melanogaster* Noppl40 variants ...... 37
2.2 The deduced amino acid sequences of DmNoppl40-RGG and DmNoppl40 ...... 38
2.3 Amino acid sequence homology of rat Noppl40, human Noppl40, frog (*X. laevis*) Noppl40, fly Noppl40 (DmNoppl40), and yeast Noppl40 (SRP40) ................. 39
2.4 Linear representations of Noppl40 homologues .......................................... 42

3.1 *In vitro* CKII and MPF phosphorylation of DmNoppl40-RGG ......................... 53

4.1 *D. melanogaster* Schneider II cells transfected to express DmNoppl40-RGG or DmNoppl40 ................................................................. 66
4.2 *D. melanogaster* Schneider II cells transfected to co-express *X. laevis* coilin and either DmNoppl40-RGG or DmNoppl40 .......................................................... 68
4.3 DmNoppl40-RGG and DmNoppl40 localize to the DFCs of *X. laevis* oocyte nucleoli ................................................................. 70
4.4 Expression of GFP-DmNoppl40-RGG and GFP-DmNoppl40 in HeLa cells ...... 72
4.5 Endogenous fibrillarin and nucleolin co-localize with DmNoppl40-RGG within nucleoli of HeLa cells ................................................................. 74
4.6 Endogenous fibrillarin, but not endogenous nucleolin co-localizes with GFP-DmNoppl40 in phase-light regions of segregated HeLa cell nucleoli ................................ 75
4.7 HeLa cells transfected to co-express GFP-DmNoppl40-RGG, GFP-DmNoppl40, or *X. laevis* coilin ................................................................. 77
4.8 Co-expression of GFP-DmNoppl40 and RFP-DmNoppl40-RGG in transfected HeLa cells ................................................................. 78
4.9 Expression of the carboxy terminal truncation, DmNoppl140ΔRGG, in transfected HeLa cells ................................................................. 80

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Abstract

The activities of non-ribosomal nucleolar proteins are now understood to be important for the normal functions of both nucleoli and Cajal Bodies (CBs). Although these proteins have been studied extensively in other eukaryotes, knowledge of non-ribosomal nucleolar proteins in Drosophila melanogaster lags far behind. The nucleolar phospho-protein of 140 kDa (Nopp140) may function to shuttle box C/D and box H/ACA small nucleolar (sno)RNAs from the nucleus to the nucleolus, where they function in the 2’-O-methylation and pseudouridylation of rRNA, respectively. Nopp140 homologues have been described in rat, human, Xenopus laevis, and yeast. This dissertation describes the cloning of cDNAs that encode two splice variants of Nopp140 in D. melanogaster. In addition, this dissertation addresses the localization patterns of the D. melanogaster Nopp140 splice variants in various cell types with respect to endogenous nucleolar proteins and CBs.

The D. melanogaster Nopp140 gene maps within 79A5 of chromosome 3. Alternative mRNA splicing yields two variants. DmNopp140 (654 residues) is the true D. melanogaster homologue of vertebrate Nopp140 in that its carboxy terminus is 58% identical to the carboxy terminus of rat Nopp140. DmNopp140-RGG (688 residues) is identical to DmNopp140 throughout its first 551 residues, but its carboxy terminus contains an extensive arginine-glycine-glycine (RGG) domain that is found in many RNA-binding proteins such as vertebrate nucleolin.

Both Drosophila Nopp140 variants localize to the dense fibrillar component (DFC) of D. melanogaster Schneider II cells and X. laevis oocytes. In HeLa cells, DmNopp140-RGG localizes to intact nucleoli, while DmNopp140 segregates nucleoli into phase-light and phase-dark regions. The phase light regions contain DmNopp140
and endogenous fibrillarin, while the phase-dark regions contain endogenous nucleolin. Both *D. melanogaster* variants co-localize to nucleoli when co-expressed in HeLa cells. Both proteins also co-localize with exogenously expressed *X. laevis* coilin to enlarged Cajal bodies (CBs) within HeLa cell nucleoli, but only DmNopp140 localizes to CBs in Schneider II cells. Both variants fail to localize to CBs in *X. laevis* oocyte nuclei. A carboxy terminal truncation, DmNopp140-ΔRGG, fails to localize to nucleoli in HeLa cells, but like DmNopp140, it localizes with exogenously expressed coilin in HeLa cell CBs.
Chapter 1

Literature Review

The nucleolus was first recognized by Fontana in 1774 (Miller, 1981). It was not until the 1960’s that the nucleolus was firmly established as the main cellular center for ribosome biogenesis in eukaryotes (Brown and Gurdon, 1964; Perry, 1962). The traditional functions of the nucleolus include the cleavage and modification of pre-rRNA precursors and incorporation of mature rRNA and ~80 ribosomal proteins into pre-ribosomal particles. Recently, new roles have been ascribed to the nucleolus. These include gene silencing, aging, and regulation of cell cycle proteins. Numerous non-ribosomal nucleolar proteins such as nucleolin/C23, fibrillarin, B23/nucleophosmin, and Nopp140 have been described over the years that play as yet undetermined roles in these various functions.

The typical metazoan interphase nucleolus consists of three morphologically distinct domains: the fibrillar center (FC), the dense fibrillar component (DFC), and the granular component (GC). The nucleolus forms primarily due to transcription of tandemly repeated rRNA genes at chromosomal loci called nucleolar organizing regions (NORs). It is now generally accepted that rRNA transcription occurs at the border between the FC and the DFC (reviewed in Scheer et al., 1999; Dundr and Raska, 1993). In mammalian cells, transcription of rDNA by RNA polymerase I initially yields a 47S primary transcript. The 47S transcript contains, in the 5’ to 3’ direction, the 5’ external transcribed spacer (ETS), the 18S rRNA, internal transcribed spacer 1 (ITS1), the 5.8S rRNA, ITS2, the 28S rRNA, and a short segment of the intergenic spacer region (IGR). The first cleavage reaction occurs very rapidly at the 5’ ETS to yield a 45S rRNA. In mammals, the 45S transcript is then further processed to yield the 18S, 5.8S, and 28S
rRNA. Following synthesis of the 5S rRNA by RNA polymerase (Pol) III outside of the nucleolus, the 5S rRNA diffuses or is carried by the La autoantigen (see below), into the nucleolus where it assembles with the 28S and 5.8S rRNAs and large subunit proteins to form large (60S) ribosomal subunits. The 18S rRNA assembles with small ribosomal proteins into small (40S) ribosomal subunits (see Fig. 1.1). After ribosomal subunit assembly, the subunits are translocated to the cytoplasm via nuclear pores. Nascent pre-rRNA transcripts are first detected in the DFC, and only later in the GC. During this vectoral movement from DFC to GC, the pre-rRNA is continually cleaved and modified until it is assembled into mature ribosomal subunits in the GC. The sequential addition of ribosomal proteins during the maturation of pre-ribosomes in mammalian cells is outlined in Figure 1.1. Processing events include cleavage of pre-rRNA into mature 18S, 5.8S, and 28S rRNA (Mougey et al., 1993; Peculis and Steitz, 1993), site specific 2’ – O – ribose methylation, and site specific pseudouridine conversion (reviewed by Lafontaine and Tollervey, 1998). Lazdins et al. (1997) used probes for various parts of the 5’ and 3’ external transcribed spacers (ETS) of mouse in conjunction with fluorescent in situ hybridization (FISH) to show that nascent pre-rRNA molecules move away from the rDNA axis in a manner consistent with this model. Also, Scheer et al. (1997) observed structures present in nucleolar cavities of grasshopper oocyte nuclei that are similar to Miller spread “Christmas trees”, which refers to the shape of transcription units with nascent RNAs radiating out from chromosomes as multiple RNA polymerase (Pol) I complexes make their way down the DNA template. These structures are enriched at the borders of the FC and DFC compartments within somatic cell nucleoli. These results strengthen the argument that rDNA transcription occurs at the FC and DFC border, and processing and ribosome assembly occur in the DFC and GC.
Figure 1.1 Maturation of the eukaryotic ribosome. Structural large and small proteins are added to 45S rRNA to form the 80S preribosome. The 55S preribosome, precursor to the 60S large ribosomal subparticle, and the 40S preribosome, precursor to the 40S small ribosomal subunit are formed from the 80S preribosome through the addition of specific L and S proteins, respectively. Adapted from Hadjiolov (1985).
Recently, a complex containing up to ten 3' - 5' exonucleases, called the exosome, has been identified in yeast. The exosome is responsible for many of the site specific cleavage reactions. All components of this complex are essential for yeast viability (Allmang et al., 1999). A human homologue of the yeast exosome, PM-Scl 100, has been localized to nucleoli in HeLa cells (Fomproix and Hernandez – Verdun, 1999).

The general structure of a nucleolus (FC, DFC, GC) is present only during interphase when rDNA transcription is at its peak. In other words, the nucleolus is “an organelle that is formed by the act of building a ribosome” (Melese and Xue, 1995). Interestingly, even though the nucleolus is considered an organelle, there is no membrane separating the nucleolus from the rest of the nucleoplasm. The nucleolus then, is a conglomeration of proteins and RNAs involved in ribosome biogenesis that come together at interphase as rRNA transcription occurs.

The metazoan nucleolus disassembles at prophase and subsequently reassembles at telophase as the two daughter cells move into interphase. The disassembly and reassembly of the nucleolus is the subject of intense research. Disassembly is initiated by the arrest of rDNA transcription (Scheer and Hock, 1999) and the suppression of rRNA processing (Dundr and Olson, 1998). Upon disassembly, the U3 small nucleolar (sno)RNA, nucleolin, and B23, all rRNA processing components, and partially processed rRNAs, are scattered into structures called nucleolus derived foci (NDF) (Dundr and Olson 1998; Pinol – Roma, 1999). This suggests that some nucleolar components are preserved upon nucleolar disassembly in prophase. Interestingly, the RNA Pol I transcriptional apparatus remains associated with nucleolar organizer regions (NORs), which are the rDNA loci on chromosomes. This is in contrast to processing components that are relocated to the NDFs or to the chromosomal periphery, or just dispersed.
throughout the cell. During telophase, pre-nucleolar bodies (PNBs) form from processing complexes in the NDFs and from the chromosome periphery. When rRNA transcription resumes in interphase, these PNBs join together at (NORs) to form a functional nucleolus. In this model, rRNA transcription drives the formation of functional nucleoli. This is not always the case, however; in early Xenopus laevis development, rDNA transcription is not needed to form nucleoli (Verheggen et al., 1998).

**RNA Species Present Within the Nucleolus**

Other species of RNA besides rRNA exist in the nucleolus. Some of these RNAs, namely the small nucleolar RNAs (snoRNAs), exist primarily to assist in the processing of rRNA (reviewed by Weinstein and Steitz, 1999). The snoRNAs constitute a class of RNAs that is absolutely critical for the correct processing of rRNA. SnoRNAs participate in both the cleavage and post-transcriptional modification of rRNA. Processing includes pseudouridine conversion and methylation of specific nucleotides within the pre-rRNA. The majority of snoRNAs, called “guide” snoRNAs, direct site specific pseudouridylation and methylation of rRNA by base pairing to the specific sites of modification. A few snoRNAs (U3 and U8) are actually involved in pre-rRNA cleavage reactions. The majority of the other snoRNAs can be divided into two groups based upon functional elements characteristic to each group. The box C/D snoRNAs contain the conserved C box (RUGAUGA; R is any purine) and D box (CUGA) elements near their 5’ and 3’ termini, respectively. Internal copies of C and D box sequences (box C’ and D’) have also been identified in some snoRNAs (Kiss-Laszlo et al., 1998). These snoRNAs form duplexes (10-21 nt) with their substrate pre-RNA site and direct the 2’ – O – methylation of the rRNA nucleotide paired to the fifth residue upstream of box D and/or D’ (see Fig. 1.2 C and D) (Smith and Steitz, 1997). Box C/D
Figure 1.2 The box H/ACA and box C/D snoRNAs.  

A The predicted structure of the hybrids between a box H/ACA snoRNA and the pre-rRNA at the sites of pseudouridine (U) formation. The sequences flanking the site of modification are base-paired to the snoRNA in a complex pseudoknot structure, while the base that is to be modified by rotation about the N3-C6 axis is free to interact with the U synthase. The base pairing involves two stretches of 3-10 nucleotides on either side of the base to be modified.  

B U formation by base rotation.  

C The predicted structure of the hybrids between a box C/D snoRNA and the pre-rRNA. The box C/D and/or C'/D'-pre-rRNA interaction generates a conserved structure with the box D or D' element placed five base pairs downstream from the site of 2'-O-methylation.  

D 2'-O-methylation of the ribose sugar moiety.  

Adapted from Weinstein and Steitz (1999).
snoRNAs associate with proteins specific for the box C/D motifs. These proteins are required for the site specific 2′ – O – methylation of rRNA (see below).

Box H/ACA snoRNAs have a conserved hairpin-hinge-hairpin-tail secondary structure that is absolutely required for pseudouridylation of rRNA. Box H (ANANNA; N is any nucleotide) lies within the hinge region while box ACA is present three nucleotides from the 3′ end of the snoRNA. Box H/ACA snoRNAs base pair with their target pre-RNA sites in two short duplexes, forming two pockets located 14-17 nucleotides upstream of box H or ACA. The target unpaired uridine is converted to pseudouridine inside of this pocket (see Fig. 1.2 A and B). Box H/ACA snoRNAs also associate with proteins that are specific only for box H/ACA motifs (see below).

Although the mechanisms of box C/D (methylation) and box H/ACA (pseudouridylation) have been intricately described, one question that remains is the purpose of the modifications. Interaction with snoRNAs may alter the overall architecture of the pre-rRNA that could be critical for achieving the conserved folds of a mature rRNA molecule. Another possibility is that modification could be important for ribosome function. Modification sites on rRNA molecules are clustered around the active site of the ribosome (Lewis and Tollervey, 2000). In certain tRNAs, pseudouridine conversion has been shown to contribute to their overall stability (Durant and Davis, 1999). Ongoing research should contribute to our knowledge regarding the functions of site specific 2′ – O – methylation and pseudouridylation in rRNA.

RNAs such as the signal recognition particle (SRP) RNA (Politz et al., 2000), and telomerase RNA (Mitchell et al., 1999), use the nucleolus for their own specific maturation (see Fig. 1.3). The signal recognition particle (SRP) is involved in protein synthesis and targeting (Politz et al, 2000). In mammalian cells the SRP consists of a 300
nt RNA, previously known as 7S RNA (Reddy et al., 1981; Walter and Johnson, 1994) and the six SRP proteins, (SRP 9, 14, 19, 54, 68, and 72). The initial function of the SRP is to bind the amino-terminal signal peptide sequence of secretory and membrane proteins through the clustered methionine residues present in SRP 54. SRP then blocks the peptidyl transferase activity of 28S rRNA, docks the polypeptide – ribosome complex at receptors on the ER, and then helps resume translation to direct the protein into membrane assembly or within the ER lumen for the secretory pathway (reviewed by Walter and Johnson, 1994). Injection of fluorescently labeled SRP RNA demonstrated that newly transcribed SRP RNA rapidly localizes to nucleoli for processing. Only then does it enter the cytoplasm (Jacobson and Pederson, 1998). GFP-tagged SRP proteins 19, 68, and 72 all localize to the nucleolus (Politz et al. 2000). Taken together, these data suggest that processing and/or assembly of SRP occurs within the nucleolus.

Telomerase is a reverse transcriptase ribonucleoprotein enzyme that synthesizes telomeric DNA repeats at the ends of eukaryotic chromosomes (see Fig. 1.3) (Blackburn and Greider, 1995). Recent studies have demonstrated that a portion of telomerase RNA co-fractionates with highly purified HeLa cell nuclei (Mitchell et al., 1999). This is consistent with the shared homology between telomerase RNA and the H/ACA box family of snoRNAs. The human telomerase RNA (hTR) contains a domain resembling a box H/ACA snoRNA. This domain is essential in vivo for hTR stability, hTR 3’ end processing, and telomerase activity (Mitchell et al., 1999).

Other examples of RNA species that are present within the nucleolus are the U6 splicesomal RNA (Tykowski et al., 1998), some tRNAs (Bertrand et al., 1998), and the snoRNAs that are the RNA components of the pre-tRNA and rRNA processing enzymes, namely RNase P and RNase MRP (see Fig. 1.3) (Lee et al., 1996). The U6 small nuclear
Figure 1.3 Overview of RNA pathways within the nucleolus. Adapted from Lewis and Tollervey (2000).
(sn)RNA is localized mainly in nuclear speckles (sites of mRNA splicing) and in Cajal bodies (see below) (Carmo-Fonseca et al., 2000; Matera and Ward, 1993). Newly transcribed U6, however, transits through the nucleolus for 2′-O-methylation (Pederson, 1998). In *Saccharomyces cerevisiae*, an isoform of isopentenyl-6-adenosine synthetase, a tRNA modification enzyme, is localized in nucleoli, suggesting that this aspect of tRNA processing occurs in nucleoli or that this enzyme modifies other non-tRNAs (Pederson, 1998). Isopentenyl-6-adenosine, however, has been reported only in tRNAs. RNase MRP plays an important role in pre-rRNA processing and consequently, most cellular RNase MRP resides within the nucleolus (Kiss et al., 1992). RNase P activity is important for tRNA processing, but recent reports suggest that RNase P may also play a role in pre-rRNA processing (Jacobson et al., 1995). When RNase P was observed in the nucleolus, many thought this reflected its role in rRNA processing. In situ hybridization studies in *S. cerevisiae*, however, demonstrate that some tRNA precursor molecules are also present within the nucleolus (Pederson, 1998).

Recently, a novel evolutionarily conserved snoRNA, U85, has been identified that directs the 2′-O-methylation of the C45 and pseudouridylation of the U46 residues in the invariant loop 1 of the human U5 splicesomal RNA. This is the first example of a snoRNA that directs modification of a snRNA and that modulates 2′-O-methylation and pseudouridylation (Jady and Kiss, 2001).

The traditional model of the nucleolus, namely a factory to produce immature ribosomal subunits, may have to be modified. Our understanding of the nucleolus now includes the assembly of and/or interaction between many other ribonucleoprotein particles that participate in gene read-out functions (Pederson, 1998). It is interesting to note that the RNA components of RNase P and RNase MRP, SRP, U6 snRNA, tRNA,
and 5S rRNA are all transcribed by RNA polymerase III. All newly transcribed RNA pol III transcripts bind to the La autoantigen, a highly conserved phosphoprotein. A fraction of the La autoantigen localizes to nucleoli, and the La autoantigen has chaperone capabilities (Wolin and Matera, 1999). An attractive model suggests that the La autoantigen binds to new RNA pol III transcripts and carries them to the nucleolus where they undergo processing and/or assembly into their respective RNP complexes (Pederson and Politz, 2000).

**Non-traditional Nucleolar Functions**

Over the past few years, several functions not traditionally associated with nucleoli have been ascribed to this sub-nuclear body. These functions include gene silencing, aging, and cell cycle regulation. Chromosomal DNA can be subdivided into two structurally distinct types, mainly euchromatin and heterochromatin. Euchromatin is generally considered to be in an “open” state and is transcriptionally active while heterochromatin is inaccessible and transcriptionally silent. In *S. cerevisiae* there are three distinct chromosome regions where gene silencing occurs: the silent mating type loci (Rine and Hersicowitz, 1987), telomeric DNA (Gottschling et al., 1990), and in the tandemly repeated rDNA genes (Gottlieb and Esposito, 1989). rDNA is highly transcribed, accounting for approximately 60% of the total RNA synthesis in active yeast cells. Interestingly, only about half of the 100-200 rDNA repeats in yeast are actively transcribed (Woolford and Warner, 1991). In fact, expression of some Pol II-transcribed genes are silenced when inserted into the rDNA repeats (Bryk et al., 1997; Fritze et al., 1997; Smith and Boeke, 1997). In all of these examples, silencing is mediated by the Sir2 protein (Gotta et al., 1997). Sir2 is a member of the Silent Information Regulatory (Sir) family of proteins that mediate gene silencing at the silent mating type loci and
telomeres. At these loci, Sir2 interacts with both Sir3 and Sir4 to mediate silencing. In the case of rDNA silencing, a new protein complex has recently been identified in budding yeast that contains Sir2, but precludes both Sir3 and Sir4. This complex, termed RENT, for regulator of nucleolar silencing and telophase exit (Shou et al., 1999; Straight et al., 1999), is connected to the yeast cell cycle machinery through the Cdc14 phosphatase (discussed below).

The RENT complex also contains the DNA-binding protein Net1. The role of Net1 in the nucleolus appears to be three-fold. First, Net1 specifically binds rDNA repeats. Net1 then recruits the Sir2 protein which mediates the silencing of some rDNA repeats. How Sir2 (and Sir3/4) silences rDNA is not known, but histones are thought to play a role in this process (Hecht et al., 1996). Second, Net1 may play a role in nucleolar integrity. In Net1Δ cells, the nucleolar protein Nop1p (homologue of vertebrate fibrillarin) redistributes over the entire nucleus, suggesting that the integrity of the nucleolus has been lost (Straight et al., 1999). Third, the nucleolus normally does not disassemble in yeast mitosis, and RENT (Net1 and Cdc14) is sequestered within the nucleolus during mitosis until telophase. At telophase, the RENT complex is disassembled, and Cdc14 is released into the nucleoplasm which allows Cdc14 to promote exit from mitosis (Shou et al., 1999). These observations suggest that the nucleolus, and more specifically the RENT complex, plays a role in gene silencing and in cell cycle control (reviewed in Visinton and Amon, 2000).

Interestingly, two other cell cycle regulators, namely p53 and MDM2, localize to the nucleolus (reviewed in Visinton and Amon, 2000). These cell cycle regulators, including Cdc14, are all involved in cell cycle checkpoint signaling. Checkpoints are surveillance mechanisms that sense intracellular stresses and halt cell cycle progression.
until the stress is removed, or until the cell is induced to undergo apoptosis. In all three of these cases, the cell cycle proteins are sequestered in the nucleolus to inhibit their activity. Some researchers liken the nucleolus to a “prison” in this respect (Straight et al., 1999). The DNA damage checkpoint causes cells to arrest in G1 or G2 of the cell cycle. In mammalian cells a key component of this pathway is the p53 protein. In response to DNA damage and/or other stresses, p53 is transiently stabilized in the nucleus where it acts as a transcription factor to upregulate genes involved in cell cycle arrest or apoptosis (Giaccia and Kastan, 1998; Prives, 1998). One particular regulation pathway of p53 involves the nucleolus. The MDM2 gene product is an inhibitor of p53 that acts by binding to p53 and preventing its transcription factor duties (Momand et al., 1992) and/or promotes p53 degradation by increasing its export to the cytoplasm (Freedman and Levine, 1998; Pomerantz et al., 1998). In response to stress, the nucleolar protein p19arf stabilizes p53 by recruiting and sequestering nuclear MDM2p to the nucleolus (Weber et al., 1999; Tao and Levine, 1999). MDM2p then sequesters or “imprisons” p53 within the nucleolus.

The last cell cycle control protein that localizes within the nucleolus is the meiosis- specific protein, PCH2 (San-Segundo and Roeder, 1999). In prophase I, homologous chromosomes undergo recombination and synapsis at a proteinaceous structure called the synaptonemal complex. Mutants defective in this process arrest due to the pachytene (a subdivision of meiosis I) checkpoint (Lydall et al., 1996). The PCH2 gene was identified in a screen for mutants defective in cell cycle arrest in response to defects in synapsis. Interestingly, the localization of PCH2p to the nucleolus requires Sir2p. These results suggest a link between the cell cycle and gene silencing machinery. Visintin and Amon (2000) find it puzzling that cells would choose sequestration in the
nucleolus as a means of inhibiting protein function. The authors suggest that the nucleolus may be the only place in the cell where a protein that functions both in the nucleus and cytoplasm can be tightly "locked away" to thus regulate its function. The nucleolus may be an excellent repository for these regulatory proteins due to its lack of a membrane. When the sequestered protein is activated, the protein would easily reach its target in the nucleus since it would not have to traverse membranes. It would reach its site of action by simple diffusion. Sequestration in the nucleolus may prove to be a common means of regulating nuclear protein function.

An intriguing connection exists between the nucleolus and aging. As yeast cells grow old they become sterile (Muller, 1985). The work of several groups showed that this sterility is caused by the loss of silencing at the mating-type loci (Smeal et al., 1996) and at telomeres (Kim et al., 1996). Kennedy et al. (1995) showed that Sir proteins (discussed above) re-localize to the "age locus" located within the nucleolus. Sinclair and Guarente (1997) suggested that redistribution of Sir proteins to the "age locus" in the nucleolus delayed aging by countering a defect occurring at this site. As described earlier, Sir proteins are responsible for silencing the rDNA genes. Consistent with this observation, uncontrolled recombination at the yeast nucleolus generates extrachromosomal circular copies of rDNA (ERCs) that accumulate within the nucleolus. The accumulation of these circles correlates with nucleolar fragmentation and cellular senescence (Sinclair and Guarente, 1997).

A further link between the nucleolus and aging occurs in patients with Werner syndrome. Patients with a mutation in the Werner gene show symptoms of accelerated aging (Epstein et al., 1966). The Werner protein is similar to DNA helicases of the Rec Q family, and it localizes in the nucleoplasm and nucleolus of human cells (Grey et al.,
Although these data are suggestive, a true link between the nucleolus and aging has yet to be established.

**Non-ribosomal Nucleolar Proteins**

Although many traditional tasks of the nucleolus (ribosome assembly) can now be described in great detail, the roles that non-ribosomal proteins play in these traditional processes and in the non-traditional processes just described are not well characterized. The most intensely studied of the non-ribosomal nucleolar proteins are fibrillarin, B23, nucleolin/C23, and Nopp140.

Fibrillarin is involved in pre-rRNA processing and ribosome assembly (Tollervey et al., 1993). Human fibrillarin (36 kDa) consists of three structural domains; an amino terminal domain rich in arginine-glycine-glycine (RGG) repeats, a central RNA binding domain (RBD), and a carboxy domain that may form alpha helices (Aris and Blobel, 1991). Unlike nucleolin (see below), the RGG domain of fibrillarin does not have a distinct function in its targeting to nucleoli. Expression of RGG deletion mutants in HeLa cells suggested that the RGG domain of human fibrillarin increases the efficiency of nucleolar targeting, but this domain is neither sufficient nor required for this targeting (Snaar et al., 2000). In contrast, the RGG domain of AtFbr1, a fibrillarin homologue in Arabidopsis thaliana, is required for its nucleolar targeting (Pih et al., 2000). Noplp, the fibrillarin homologue in budding yeast, is an essential component of box C/D snoRNPs that participate in the 2′-O-methylation of rRNA. In fact, a recent report suggests that fibrillarin from the hypothermophile, Methanococcus jannaschii, may function as the methyltransferase itself (Wang et al., 2000). Although fibrillarin is a component of Cajal bodies (see below), its highest steady state concentration resides within the DFC of nucleoli.
B23/numatrin/nucleophosmin/NO38 is a multifunctional protein. B23 is often referred to as a putative ribosome assembly factor. Like other nucleolar proteins (C23, Nopp140) it contains an alternating acidic/basic region. B23 is a substrate for casein kinase II (CKII) (Chang et al., 1990) and Mitosis-Promoting Factor (MPF) (Peter et al., 1990) (see below). It also binds nuclear and nucleolar proteins, and engages in nucleocytoplasmic shuttling (Valdez et al., 1994; Dundr et al., 1996). After a cell enters mitosis and the nucleolus disassembles, B23 redistributes to cytoplasmic bodies and to PNBs (Ochs et al., 1985; Dundr et al., 1997; Zatsepina et al., 1997). B23 also redistributes to the mitotic spindle poles and interacts with the nuclear matrix protein, NuMA, during mitosis (Zatsepina et al., 1999). This redistribution of B23 to spindle poles (centrosomes) appears to be dependent upon B23's hyperphosphorylation. It also appears that B23 may play an important role in the regulation of centrosome duplication. B23 dissociates from centrosomes when phosphorylated by CDK2/cyclin E at late G1 of the cell cycle, which then may trigger initiation of centrosome duplication. During mitosis, B23 is dispersed to NDFs, the cytoplasm, and to perichromosomal regions. Upon nuclear reassembly, some B23 remains associated with centrosomes. After mitosis, each daughter cell receives one centrosome containing B23 until the next cell cycle when it is phosphorylated by CDK2/cyclin E (Okuda et al., 2000).

Nucleolin/C23 has been implicated in chromatin structure, rDNA transcription, rRNA maturation, ribosome assembly, and nucleocytoplasmic transport (reviewed in Ginisty et al., 1999). Nucleolin (110kDa) is a modular protein. Its amino terminal domain contains alternating acidic/basic regions. This is followed by four consensus RBDs, and a carboxy-terminal RGG domain (Lapeyre et al., 1987). Ghisolfi et al. (1992b) showed that the RGG domain of hamster nucleolin interacts non-specifically.
with RNA. The RGG domain of nucleolin may facilitate the interaction of the upstream RBDs with RNA (Heine et al., 1993), specifically rRNA (Ghisolfi et al., 1992a). On the other hand, the RGG domain of the nucleolar phosphoprotein Nopp 44/46 of the protozoan parasite *Trypanosoma brucei* was both necessary and sufficient for interactions with nucleic acids *in vitro*, even though Nopp 44/46 contains no RBDs (Das et al., 1998). This contrasts with the case of nucleolin in which nucleolin’s RGG domain serves to facilitate interaction with RNA through the RBD (see above). The possibility remains that the large number of RGG motifs (22-26) in Nopp 44/46 allow the protein to sufficiently interact with RNA without the presence of an RBD.

Alternatively, RGG domains may function in protein-protein interactions. For example, the RGG domain of hnRNPA1 interacts with itself and with other hnRNP proteins (Cartegni et al., 1996). The RGG domain of mouse nucleolin and the fission yeast homologue of nucleolin, GAR2, interacts with specific ribosomal proteins (Bouvet et al., 1998; Sicard et al., 1998). These results suggest that the RGG domain can facilitate protein/protein interactions. Further, these interactions can be quite specific in that mouse nucleolin appears to interact with only a subset of ribosomal proteins (Bouvet et al., 1998).

**Nopp140**

Meier and Blobel (1990) discovered a protein of 140 kDa (p140) in an assay to look for nuclear proteins that interact with the nuclear localization signal (NLS) of the SV40 large T antigen. Antibodies raised against p140 localized this protein within the nucleolus of Buffalo rat liver (BRL) cells. Monoclonal antibodies directed against a yeast homologue of fibrillarin (Nop1p) co-localized with the p140 antibodies, suggesting that p140 was present specifically in the DFC. Cloning and sequencing of a cDNA that
encoded p140 revealed a unique peptide sequence. Two thirds of the protein consists of only five amino acids, namely serine (17.2%), lysine (16.1%), alanine (14.3%), glutamic acid (9.7%), and proline (9.2%). Most of the serines and acidic amino acids are clustered into a central domain of ten acidic repeats separated by basic amino acid stretches (Meier and Blobel, 1992). The conserved amino acids within the acidic repeats formed several consensus casein kinase II (CKII) phosphorylation sites (Marin et al., 1986; Kuenzel et al., 1987). Specifically, forty-five serines within these acidic repeats are the direct phosphorylation substrate sites for CKII. Following phosphorylation of these initial forty-five serines, the remaining serine residues within the acidic repeats also become acceptor sites for CKII phosphorylation. Potentially, this leads to stretches of 13-17 negatively charged amino acid residues in a row.

P140 was renamed Nopp140 for nucleolar phosphoprotein of 140 kDa due to the protein's requirement of phosphorylation to bind positively charged NLS containing peptides. Meier and Blobel (1992) showed that Nopp140 shuttles between the nucleus and the cytoplasm. Antibodies raised against Nopp140 and tagged with rhodamine were injected into living cells, and subsequent incubation demonstrated that the Nopp140 antibodies accumulated within the DFC of nucleoli of BRL cells (Meier and Blobel, 1992). Further, immunogold labeling of thin sections of BRL cells showed DFC labeling. Interestingly, the gold particles labeled curvilinear tracks that ran from the DFC, through the nucleoplasm, and to nuclear pore complexes (NPCs). Taken together, these results suggest that Nopp140 shuttles on these tracks, perhaps to import ribosomal proteins or export ribosomal subunits (Meier and Blobel, 1992).

Pai et al. (1995) discovered the human homologue of Nopp140, p130, in a screen using monoclonal antibodies to search for proteins whose concentration oscillates in the
cell cycle. This group led by Ning-Hsing Yeh went on to show that p130 is hyper-phosphorylated in mitosis as compared to p130 present during interphase, and that this hyperphosphorylation led to its dispersal into the mitotic cytoplasm. Maturation promoting factor (MPF) (see Chapter 3) was implicated in p130’s phosphorylation. Yeh’s group also showed that at the end of mitosis, p130 aggregated in structures reminiscent of PNBs before moving into intact nucleoli in early G1. This is not surprising since several non-ribosomal nucleolar proteins are observed in PNBs during nucleogenesis (reviewed in Pinol-Roma, 1999; Dundr et al., 2000).

Yeh’s group also showed that human Nopp140 (p130) was present in a complex that included RPA194 (the largest subunit of RNA Pol I). Immuno-fluorescence co-localizations of p130 and RNA Pol I in HeLa cells suggest that this interaction may also occur in vivo. Overexpression of full-length p130 or specific carboxy terminal truncations of the protein resulted in mislocalization of RNA Pol I. From these results, they suggested that p130 plays a role in rDNA transcription and/or the maintenance of nucleolar integrity. Other studies suggest that Nopp140 can interact with the alpha-1 acid glycoprotein/enhancer binding protein (AGP/EBP, a transcription factor) and TFIIB to synergistically activate the AGP gene (Miau et al., 1997).

Finally, casein kinase II (CKII) specifically interacts with Nopp140 (Li et al., 1997). This observation was expected because of the many putative CKII phosphorylation sites located within the acidic repeats of Nopp140. CKII is regulated by signal transduction. Therefore, the association between Nopp140 and CKII suggests that Nopp140 is a terminal acceptor in signal transduction phosphorylation cascades that could ultimately modulate its function and thus nucleolar function.
The Nucleolus and Cajal Bodies

The nucleolus has several “nuclear neighbors” or nuclear subdomains that physically interact with the nucleolus. Some of these neighbors are the perinuclear compartment (Matera et al., 1995; reviewed in Huang, 2000), the hnRNP L domain (Pinol-Roma et al., 1989), the Oct 1/PTF/transcription (OPT) domain (Grande et al., 1997; Pombo et al., 1998), and Cajal bodies.

Perinuclear components are observed primarily in transformed cells rather than normal primary cells in culture (Huang, 1997). The functions of the perinuclear compartment are unknown. Perinuclear compartments could arise due to the up-regulation of specific genes in certain cells (cancer cells). Perinuclear compartments may preferentially associate with these specific genes, much like Cajal bodies associate with histone gene clusters (see below). These structures, however, are enriched in small RNAs transcribed by RNA pol III (RNase MRP RNA, RNase P RNA, and hY RNAs) (Matera et al., 1995), the polypyrimidine tract binding (PTB) protein (Ghetti et al., 1992), the CUG binding protein (Timchenko et al., 1996), and the KH-type splicing regulatory protein.

The hnRNP L domain is so named because of its high concentration in the hnRNP L protein. Although the function of the hnRNP L domain is unknown, the hnRNP L protein is known to stabilize specific mRNAs (Shih and Claffey, 1999). The OPT domain preferentially associates with chromosomes 2, 6, and 7, and it is enriched in the transcription factors Oct1 and PTF. The OPT domain also contains RNA polymerase II and III transcription sites (Grande et al., 1997; Pombo et al., 1998). The function of the OPT domain is also unknown. Of these nuclear subdomains, the Cajal body is the best characterized (reviewed by Gall, 2000), and most related to this work.
Cajal Bodies

The Cajal body was first identified by Santiago Ramon y Cajal in 1903 (Cajal, 1903) as an accessory body to nucleoli in neuronal cells. The Cajal body was “rediscovered” in non-neuronal cells by Monneron and Bernhardt (1969). These authors coined the name “coiled body” due to the structure’s coiled appearance by electron microscopy of thin sections of mouse, rat, and human cells. In the same year, Hardin et al. (1969) published similar images of coiled bodies in nuclei of rat trigeminal neurons. Over the past few years there has been a virtual explosion of coiled body-based research (reviewed by Gall, 2000). Joe Gall has re-named coiled bodies, Cajal bodies, in honor of Ramon y Cajal (Gall et al., 1999). Thus, for the rest of this dissertation, the term Cajal bodies will be used to describe coiled bodies.

Cajal body research changed dramatically when an autoimmune serum that labeled “dots” in interphase nuclei was discovered by Raska et al. (1991). These dots were identified as Cajal bodies by immunogold labeling of thin sections. With this serum in hand, researchers now had a tool to identify Cajal bodies and Cajal body components in many different tissues. This serum recognized a single protein of 80 kDa on western blots. The protein was subsequently named p80 coilin.

Although p80 coilin has been generally accepted as a marker for Cajal bodies, most coilin is dispersed throughout the nucleoplasm, and it is enriched in regions associated with interchromatin granule clusters (IGCs) (Puvion-Dutilleul et al., 1995). Cajal bodies contain the splicing proteins CPSF100 and CstF77, along with the SR proteins. Transient expression of GFP-tagged snRNP Sm proteins in mammalian cells shows that the Sm proteins first concentrate in Cajal bodies before moving to nuclear speckles or IGCs, the sites of pre-mRNA processing (Sleeman and Lamond, 1995).
Cajal bodies contain splicing snRNAs and snoRNAs. Injection of fluorescently labeled U3 and U8 snoRNAs into *X. laevis* oocytes results in their accumulation in Cajal bodies before relocating to nucleoli (Narayanan et al., 1999). The results of these experiments suggest a role for Cajal bodies in the transport and/or maturation of both the snRNAs necessary for splicing and the previously mentioned snoRNAs necessary for pre-rRNA processing. Although Cajal bodies contain high concentrations of many snRNAs and snoRNAs, no target pre-messenger RNAs or pre-ribosomal RNAs are present within Cajal bodies, suggesting that splicing and processing do not occur within these bodies.

RNA polymerases I, II, and III, along with TFIIA and TFIIF (RAP74) also reside in Cajal bodies, at least in *X. laevis* oocytes. Gall et al. (1999) propose that Cajal bodies may function as assembly sites for major transcription machines or “transcriptosomes”, at least in *X. laevis* oocytes. Gall et al. (1999) speculate that the three eukaryotic RNA polymerases associate with their respective transcription and processing factors in Cajal bodies, from which they are transported as preassembled complexes (transcriptosomes) to the chromosomes, the actual sites of transcription. Transport of transcriptosomes would be carried out by B-snurposomes, which in the *X. laevis* oocyte correspond to speckles or pre-mRNA splicing factor domains in somatic cells. Cajal bodies, or C-snurposomes in the oocyte, preferentially co-localize with specific gene loci such as histone gene clusters (Callan et al., 1991) and the U1, U2, and U3 snRNA gene loci (Smith et al., 1995). Association of Cajal bodies with the U2 gene locus depends upon U2 transcription at this site (Frey et al., 1995), thus suggesting a role for Cajal bodies in the delivery of processing factors to new transcripts.
There is a clear relationship between the nucleolus and Cajal bodies. Expression of mutant forms of coilin disrupt normal nucleolar and Cajal body structure (Bohmann et al., 1995b), and expression of Nopp140 mutants leads to changes in Cajal body structure (Isaac et al., 1998). Also, exposure of mammalian cells to okadaic acid, a serine/threonine protein phosphatase inhibitor, results in the accumulation of p80 coilin and splicing snRNPs within nucleoli (Lyon et al., 1997). Obviously there is interplay between the nucleolus and Cajal bodies. Both structures are present in many cell types. In some cells, however, a visible Cajal body is not necessary for viability. Microinjection of anti-coilin antibodies leads to a loss of visible CBs, but the cells lacking CBs are viable for mRNA splicing (Almeida et al., 1998).

Cajal bodies are usually 0.5 – 1.0 μm in diameter, but they vary in size and number depending upon the cell type. Like nucleoli, Cajal bodies disassemble during mitosis and reform in interphase at early to mid-G1. P80 coilin is a phosphoprotein, and phosphorylation of this protein may play a role in the assembly/disassembly cycle of Cajal bodies (Carmo-Fonseca et al., 1993). The kinase, CDK2 – cyclin E, is enriched in Cajal bodies during the G1-to-S phase transition of the cell, thus suggesting that CDK-cyclin E is responsible for the phosphorylation of p80 coilin (Liu et al., 2000). This phosphorylation may facilitate the breakdown of Cajal bodies at mitosis.

Platani et al. (2000) recently used a stable HeLa cell line that constitutively expresses a GFP-tagged version of p80 coilin to shed new light on Cajal body dynamics. Their results suggest that Cajal bodies are extremely mobile, and that there are two classes of Cajal body size, mainly >0.4 μm and <0.2 μm. Their data also suggests that the concentration of p80 coilin in Cajal bodies is in constant flux. The constant
movement and flux of specific antigens to and from Cajal bodies suggest that Cajal bodies are dynamic structures that may have multiple functions (see above).

**Nopp140 and Cajal Bodies**

The focus of Nopp140 research shifted when Isaac et al. (1998) discovered that Nopp140 was a component of Cajal bodies as well as nucleoli. At that time, investigation focused on the possible roles of Nopp140 in rDNA transcription and ribosomal subunit shuttling. With the discovery of Nopp140 in Cajal bodies, however, researchers now had to search for Nopp140 roles that were common to both nucleoli and Cajal bodies. Isaac et al. (1998) prepared and expressed mutants of Nopp140 in HeLa cells. Exogenous expression of the conserved carboxy terminus of Nopp140 chased endogenous Nopp140, fibrillarin, and NAPS7 (all common to both Cajal bodies and nucleoli) out of nucleoli. A specific interaction between Nopp140 and the amino terminus of p80 coilin was observed through yeast two hybrid analysis and co-immunoprecipitation (Isaac et al., 1998). The interaction of the amino terminus of p80 coilin and Nopp140 was lost as soon as Nopp140 was truncated for any of its domains, suggesting that full length Nopp140 was required for this interaction. This interaction between p80 coilin and Nopp140 was very gratifying because Cajal bodies were traditionally believed to have a physical connection with nucleoli (Cajal, 1903). Now for the first time, Nopp140 provided a molecular link between these two nuclear bodies (as will be discussed further in Chapter 4).

A direct interaction between Nopp140 and NAP57 (Nopp140 associated protein of 57 kDa) was also established (Meier and Blobel, 1994). NAP57 shares homology with yeast Cbf5p (Cadwell et al., 1997), with the product of the *Drosophila Nop60B* gene (Phillips et al., 1998), and with dyskerin, the product of the human *DKC1* gene. Mutation
of *DKC1* causes a bone marrow failure, a disorder called dyskeratosis congenita (Heiss et al., 1998). Interestingly, all of these proteins share homology with TruB, a pseudouridine synthase for tRNAs in *E. coli*. Thus, it was not surprising when Nopp140 was discovered to interact with mammalian box H/ACA snoRNPs (Yang et al., 2000). Box H/ACA snoRNPs direct the conversion of uridine to pseudouridine in pre-rRNA (see above). As expected, NAP57 co-localizes with Nopp140 in the DFC of nucleoli and in Cajal bodies (Meier and Blobel, 1994).

Interactions between Nopp140 and box H/ACA and box C/D snoRNPs are most likely mediated by NAP57 (see above) and NAP65 respectively. NAP57 is a protein specific for box H/ACA snoRNPs, while NAP65 is specific for box C/D snoRNPs (see Fig. 1.4). NAP65 homologues have been demonstrated in most species, including plants, but only the yeast NAP65 homologue, Nop5/58p, has been characterized to date (Gautier et al., 1997; Wu et al., 1998). Nop5/58p is an integral part of box C/D snoRNPs (Lafontaine and Tollervey, 1999). Further, the human homologue of NAP65, hNop5/Nop58 is also a component of box C/D snoRNPs (Lyman et al., 1999).

The conserved carboxy terminus of Nopp140, when overexpressed on its own, caused dominant-negative effects in HeLa cell nucleoli. Specifically, it depleted snoRNP proteins from HeLa cell nucleoli (Yang et al., 2000). From these results, the authors proposed that Nopp140 functions as a chaperone for these snoRNPs. Box C/D and H/ACA snoRNPs function to modify rRNA in the nucleolus, but they are synthesized in the nucleoplasm. The box C/D element of the snoRNA is required for nucleolar localization (Samarsky et al., 1998), and box H/ACA sequence elements probably function in the same manner (Bortolin et al., 1999). Whether or not extrinsic transport factors or certain snoRNP proteins function to shuttle snoRNAs between the nucleoplasm
and nucleolus remains unknown. Nopp140 is a shuttling protein, and it is a component of box C/D and box H/ACA snoRNPs. Thus, Nopp140 is a candidate for a snoRNP chaperone. This theory is supported in yeast, where genetic depletion of Srp40p the yeast Nopp140 homologue, in a conditional lethal strain, specifically reduces the level of box H/ACA snoRNAs within nucleoli (Meier, 1996).

Nopp140 homologues are present in rat, humans (p130), and *Xenopus* (Nopp 180) (Cairns and McStay, 1995). Treacle is a human protein related to Nopp140 in that it contains the signature central domain consisting of alternating acidic and basic regions (Dixon et al., 1996, 1997; Wise et al., 1997). Like Nopp140, treacle is highly phosphorylated by CKII. Treacle also localizes to the DFCs of nucleoli, but it does not localize to Cajal bodies (Isaac et al., 2000). Perhaps this is due to differences in the proteins at their amino and/or carboxyl terminal ends. Mutations in the treacle gene, *TCOFI*, cause the Treacher Collins syndrome, an autosomal dominant disorder that affects craniofacial development (Dixon, 1996). Because treacle exhibits primary amino acid structure similar to that of Nopp140 and localizes to the nucleolus, it may perform functions related to ribosome assembly. Mutations in treacle could have far reaching effects during development, when high levels of protein synthesis are critical.

With Nopp140 having only been discovered in 1990, research regarding Nopp140's role in the processes related to nucleoli and Cajal bodies is still in its infancy (Meier and Blobel, 1990). Nopp140 is unique in that it does not contain any RNA binding domains (RBDs) like most other non-ribosomal nucleolar proteins. If Nopp140 does not bind RNA, what role does it play in the RNP rich nucleolus and Cajal bodies? The only clues about Nopp140 function come from its associated proteins, namely p80 coilin, CKII, NAP57, and NAP65. Since Nopp140 shuttles between the nucleolus and
Figure 1.4 Interactions of Nopp140 with box H/ACA and box C/D snoRNPs. The names of the mammalian proteins are printed in black, and the names of the homologous yeast proteins are printed in white. Box H/ACA and box C/D snoRNPs contain different subsets of snoRNAs and proteins. Nopp140, however, interacts with both types of snoRNPs, possibly as a chaperone. The interaction of Nopp140 with box H/ACA snoRNPs is greater than its interaction with box C/D snoRNPs (thickness of arrows). Nopp140 also interacts with CKII (Li et al., 1997) and with p80 coilin (Isaac et al., 1998). Adapted from Isaac et al. (1998).
cytoplasm, a current model suggests that Nopp140 functions as a chaperone for snoRNPs in both yeast and vertebrate cells (Yang et al, 2000). Nopp140 could also function in the biogenesis of snoRNPs, or perhaps in their nucleolar function. Nopp140 therefore, could serve multiple roles and overlap in all of these functions.

This dissertation describes the isolation of two splice variants of Nopp140 in *Drosophila melanogaster*. One splice variant, DmNopp140, is the *bona fide* homologue of vertebrate Nopp140 in that it contains a central domain of alternating acidic and basic regions and an evolutionarily conserved carboxy terminus. The other splice variant, DmNopp140-RGG, is identical to DmNopp140 except that the evolutionarily conserved tail is replaced by a prominent tail rich in arginine-glycine-glycine repeats. This dissertation describes the isolation, sequencing, and cloning of cDNAs corresponding to these splice variants. In addition, this dissertation demonstrates the phosphorylation of DmNopp140 by CKII and by MPF. It also demonstrates that the differences in the carboxy terminus of the two proteins lead to differences in localization patterns in Schneider II cells, and HeLa cells. Exogenous expression of DmNopp140 in HeLa cells causes their nucleoli to segregate into phase-light and phase-dark regions. The segregation of nucleoli caused by DmNopp140 permits future molecular characterizations of the nucleolus. Herein are presented the first discussions of the Nopp140 splice variants in *D. melanogaster*. It is hoped that these results and future molecular-genetic analysis will further the understanding of Nopp140 proteins and non-ribosomal nucleolar proteins in general.
Chapter 2

Identification of Two Splice Variants of Nopp140 in *Drosophila melanogaster*: Screening, Sequencing, and Comparison of Nopp140 Homologues Across Species

Introduction

The traditional functions of the nucleolus include the transcription of rDNA, and the modification and packaging of rRNA into pre-ribosomal particles (reviewed by Busch and Smetana, 1970; Hadjiolov, 1985). Recently, functions not typically associated with nucleoli have been ascribed to these nuclear bodies. These include gene silencing, assembly of signal recognition particles (SRPs), cell cycle regulation, and aging (reviewed by Pederson, 1998; Pederson and Politz, 2000; Visinton and Amon, 2000; Carmo-Fonseca, 2000). Several non-ribosomal nucleolar proteins whose presence is required for the normal operations of nucleoli have been described in yeast and metazoans. They include B23, nucleolin, fibrillarin, and Nopp140 (see Chapter 1).

Precedent has been set for naturally occurring splice variants of mRNAs encoding nucleolar proteins. Two splice variants of B23 occur in rat (Chang and Olson, 1989, 1990; Wang et al., 1994). B23.1 is the prominent nucleolar protein expressed in all tissues, while B23.2 is a shorter variant that localizes to the cytoplasm and perhaps to the nucleoplasm, but not to the nucleolus (Wang et al., 1993). Alternative splicing of the B23 pre-mRNA results in the deletion of the carboxy terminal 35 amino acid residues and the substitution of two additional upstream residues to convert B23.1 to B23.2 (Chang and Olson, 1989, 1990). *Xenopus laevis* expresses two versions of nucleolin that are probably encoded by separate genes (Meßmer and Dreyer, 1993). *X. laevis* is pseudo-tetraploid (Kobel and DuPasquier, 1986), and the two nucleolin genes may have descended from a common ancestral gene by duplication and then divergence.
Interestingly, two isoforms of human Nopp140 (p130) have been reported (Pai and Yeh, 1996). P130α is the predominant isoform, while p130β is present to a much lesser degree. P130β contains an insert of 10 amino acids within a region corresponding to the fourth proline-rich basic domain of p130α. Pai and Yeh (1996) concluded that the 10 amino acid insert has no discernable effect on the interactions of p130 with other nucleolar components in interphase cells.

A homologue of Nopp140 in yeast (S. cerevisiae) is the product of the SRP40 gene (Meier, 1996). SRP40 is a nucleolar protein of 41kDa that is much like Nopp140 in that it consists of 48% serine residues that are clustered within two long acidic stretches. These acidic stretches contain 52 consensus CKII phosphorylation sites (see Chapter 3). In addition, the carboxy-terminal 51 amino acid residues of SRP40 exhibit 59% sequence identity to the last 51 amino acid residues of rat Nopp140. SRP40 and rat Nopp140 are immunologically related in that antibodies directed against rat Nopp140 recognize SRP40 on western blots (Meier, 1996).

While Nopp140 and SRP40 are similar in many ways, they also differ. First, Nopp140 is phosphorylated to a much higher degree than SRP40. SRP40 is already a very acidic protein and its phosphorylation does not introduce any more negative charges, despite the 52 consensus CKII phosphorylation sites present in the protein. Meier and Blobel (1992) showed that phosphorylation of rat Nopp140 by CKII was required for binding NLS-containing peptides, probably by introducing many negative charges. Thus, SRP40 is a constitutively acidic protein that does not contain this level of regulation found in rat Nopp140. Second, SRP40 lacks consensus MPF phosphorylation sites (see Chapter 3).
Yeast is an excellent model organism, but the fact remains that the yeast nucleolus is different than that of metazoans. The yeast nucleolus is a crescent shaped structure that does not disassemble during mitosis like the nucleolus of metazoans. Thus, all information gathered from studying nucleolar proteins in \textit{D. melanogaster} is applicable to higher organisms, including humans.

\textit{D. melanogaster} is an excellent model organism in which to investigate the molecular genetics and molecular cytology of nucleolar proteins. The \textit{D. melanogaster} egg chamber is ideal for studying the nucleolus due to the tremendous production of ribosomes throughout oogenesis. The \textit{D. melanogaster} developmental cycle has been extensively characterized and many techniques already exist to study the embryo and larval stages of this organism. The life cycle of \textit{D. melanogaster} is rapid which makes it an excellent organism to explore the link between the nucleolus and aging (see Chapter 1).

Finally, much remains unknown about nucleolar proteins in \textit{D. melanogaster} as compared to other organisms. The only \textit{D. melanogaster} nucleolar specific proteins reported to date are Nop60B (Phillips et al., 1998), which is the homologue of yeast Cbf5p (a putative pseudouridine synthase) (Lafontaine et al., 1998), the \textit{D. melanogaster} modulo gene product (Perrin et al., 1999), which modulates chromatin structure, and the \textit{Dnop5} gene product (Vorbruggen et al., 2000), which is a conserved member of the Nop/Sik family of the conserved rRNA processing factors. This chapter describes the isolation and cloning of cDNAs corresponding to two splice variants of Nopp140 in \textit{D. melanogaster}. The protein sequences of the splice variants are compared to Nopp140 homologues from rat, human, frog, and yeast.
Materials and Methods

Standard molecular biology techniques were used to screen a D. melanogaster stage 10 egg chamber cDNA library with a random primed [32P]-labeled sub-clone of a X. laevis nucleolin cDNA (Rankin et al., 1993). The subclone was a region from a Nse I site at the translation start to a Pst I site 444 bp downstream. This 444 bp subclone encodes most of the N-terminal alternating acidic/basic region present in X. laevis nucleolin. Low stringency washes were used (2 x SSC without SDS at room temperature) to detect several strongly positive plaques, four of which were picked and re-screened to establish clonal purity. These lambda clones were amplified and their genome was prepared and then digested with Eco R1 to remove the D. melanogaster cDNA inserts from the left and right arms of the lambda phage genome. The D. melanogaster cDNAs were ligated into pBluescript KS+ (Strategene, La Jolla, CA).

Two of the D. melanogaster cDNAs, B72A (787 bp) and B71 (1300 bp) were sequenced in both directions using Sangers’ dideoxy method for DNA sequencing. The Sequenase quick denature kit (USB, Cleveland, OH) was used for all sequencing reactions according to the manufacture’s instructions. The deduced translation product of these two overlapping clones constituted an alternating acidic/basic domain similar to the 5’ X. laevis nucleolin subclone originally used to screen the library.

The B72A insert was used to re-screen the D. melanogaster library under increased stringency (0.5 x SSC, 0.1 x SDS at 60° C). Rescreening identified many larger cDNA inserts that were subsequently subcloned into pBluescript KS+. Primers designed to sequence the B72A and B71 clones were used to determine if these newly identified D. melanogaster clones contained overlapping sequence. Sequence analysis using the most 5’ primer to date (called 5’FINISH, see Appendix A), showed that clone
F2 (2400 bp) contained the 5’ end of a Nopp140-like cDNA including the translation start site. Upstream (55 bp) of the translation start site, however, was a cloning artifact that consisted of sequence from a centrosome protein (acc. no. Z50021). Possibly, this sequence was an artifact from the construction of the original cDNA library (see above).

The F2 clone, however, was not a full-length cDNA; it did not contain the 3’ end of the cDNA (as determined by screening with 3’ most primer at the time, 3’NOPP). In an effort to search for a full-length cDNA or the 3’ end of the cDNA, more lambda clones were prepared and cloned into pBluescript KS+. The cDNA clones were screened with 3’Nopp, and one clone, E4 (1200 bp), contained sequence downstream of any previously determined sequence. Another primer, (3’NOPPA) was subsequently designed from the newly acquired sequence. Sequencing from this primer proved difficult, however, because of the many stop bands and compressions that were continuously encountered. To sequence past this GC rich region of the gene, sequencing reactions were performed as normal, except that all steps were performed at 50°C instead of the recommended 37°C (suggestion from Sequenase version 5.0 manual, USB). Fortunately, enough sequence was determined with this method to design another primer, 3’NOPPB, that was further downstream than 3’NOPPA.

The sequence already in hand was used to query the database at the Berkely Drosophila Genome Project (BDGP) (www.fruitfly.org). The query sequence matched perfectly to several expressed sequence tag (EST) clones synthesized by the BDGP staff. Because only the sequence from the 5’ ends of the EST clones was available (average of 500 bp), clone LD10913 whose 5’ terminal sequence matched that of the translation start region of the query sequence was chosen. LD10913, a potential full-length cDNA...
encoding a *D. melanogaster* homologue of vertebrate Nopp140 was ordered from Genome Systems, Inc. (St. Louis, MO).

The LD10913 cDNA insert was removed from its vector by *Eco* RI and *Xho* I digestion. Restriction analysis of just the insert showed that LD10913 was roughly 3600 bp in length. LD10913 comes from the LD library (made by Lin Hong) made from mRNAs isolated from 0-22 hr *D. melanogaster* embryos. The cDNAs were made using Stratagene's ZAP-cDNA synthesis kit. A series of primers were used to obtain the 3' sequence of LD10913 (primers were 3'NOPPA-E, T3). Sequencing of the 5' half of LD10913 was also performed using primers complimentary to the *D. melanogaster* clones obtained from the stage 10-egg chamber library. Thus, LD10913 turned out to be a complete cDNA encoding a *D. melanogaster* Nopp140-like protein.

DmNopp140-RGG was then used to probe the BDGP database. The gene was found in 79A5. This region is proximal to the centromere on polytene chromosome 3L. Interestingly, the BDGP database predicted that the gene encodes two transcripts, CT22833 and CT22845. The translation product of CT22845 matches that of DmNopp140-RGG, and its sequence is nearly identical to the cDNA sequence encoded by LD10913 (acc. no. AF162774). In an effort to obtain a cDNA encoding the true *D. melanogaster* homologue of Nopp140, two EST clones, LD33426 and SD10348, were ordered from Research Genetics (Huntsville, AL). LD33426 was subcloned differently in only one way. LD clones with identification numbers greater than 21101 were directionally cloned into the plasmid pOT2 rather than pBluescript. Another library, the SD library, was made by Ling Hong from mRNAs isolated from *D. melanogaster* Schneider II cells.
To determine the identity of the cDNAs (DmNopp140 vs. DmNopp140-RGG), LD33426 and SD10348 were subjected to restriction digestion analysis. Inserts were cut from the pOT2 vector using Eco R1 and Xho I, and resolved on a 0.7% agarose gel. The cDNA inserts were isolated from the gel using standard molecular biology techniques. The inserts were subsequently digested with either Pvu I or Ava I to determine their identity. According to predicted restriction digestion maps, Pvu I should cut the DmNopp140-RGG cDNA, but not the DmNopp140 cDNA, while Ava I should cut the DmNopp140 cDNA, but not the DmNopp140-RGG cDNA.

Results

In an original effort to recover nucleolin or Nopp140 proteins from D. melanogaster, a 5' subclone of X. laevis nucleolin (Rankin et al., 1993) was used to probe a cDNA library prepared from stage 10 egg chambers. This subclone corresponded to the alternating acidic/basic domain in the N-terminus of X. laevis nucleolin. Translation of an open reading frame (ORF) corresponding to two positive overlapping clones (B72A and B71) resulted in a peptide sequence of repeating acidic/basic domains. The acidic domains are rich in serine, aspartic acid, and glutamic acid, while the basic domains are rich in lysine, alanine, and proline. The acidic domains are similar in sequence to those in nucleolin, and they have been implicated in the binding of NLS-containing proteins (reviewed in Xue and Melese, 1994).

After sequencing a cDNA clone (F2) that encodes the amino terminus, a BLAST search against the BDGP database was performed. The 5' sequence of F2 corresponds to the 5' sequence of LD10913. LD10913 was later determined to be a full length cDNA that encoded DmNopp140-RGG (acc no. AF162774). DmNopp140-RGG resembles vertebrate Nopp140 in that it contains alternating acidic and basic domains. The D. melanogaster.
*melanogaster* protein differs though in its carboxy terminal end where it contains an extensive arginine/glycine rich tail similar to the carboxy RGG domain in vertebrate nucleolin (Lapeyre et al., 1987). As discussed previously (Chapter 2), GENIE and GENESCAN programs predicted two transcripts CT22822 (DmNopp140) and CT22845 (DmNopp140-RGG) from the *D. melanogaster* Nopp140 gene (CG7421). With the DmNopp140-RGG cDNA in hand, we knew that this variant was indeed expressed. To determine if the DmNopp140 variant (CT22833) was expressed, several EST clones were ordered from Research Genetics (Huntsville, AL) for the purpose of restriction digest screening. Using the predicted restriction maps of both variants, it was determined that the DmNopp140 variant was also expressed (Fig. 2.1).

The DmNopp140-RGG sequence was then used to probe the BDGP database. The *D. melanogaster* Nopp140 gene maps within 79A5 of the proximal left arm of polytene chromosome 3 (see http://hedgehog.lbl.gov:8000/cgi-bin/annot/gene?CG7421). The *D. melanogaster* Nopp140 gene was designated CG7421 (conceptual gene no. 7421) by the BDGP. Some of the genes flanking *D. melanogaster* Nopp140 have been characterized previously. *Eagle*, located in 79A4 is a gene that encodes a steroid hormone receptor/transcription factor. The gene encoding cyclin H is also located nearby.

CG7407 (unknown function), CG7414 (unknown function), CG7148 (unknown function), and CG7145 (a 1-pyrroline-5-carboxylate dehydrogenase-like enzyme) are nearby genes, that like *D. melanogaster* Nopp140, were identified by Genie and Genescan (see below). CG7145 contains an enhancer-promoter type P-element transposon (EP(3)3138) within its promoter region. EP(3)3138 is located approximately...
Figure 2.1 Expressed sequence tags exist for both *D. melanogaster* Nopp140 variants. Digestion patterns were predicted from restriction maps to verify the clones. The purified cDNA insert that encodes conceptual transcript (CT) 22833 (DmNopp140) was left undigested (lane 1), incubated with *Pvu* I that failed to cut as predicted (lane 2), or digested with *Ava* I that generated two bands of predicted size (lane 3). The purified cDNA insert that encodes conceptual transcript 22845 (DmNopp140-RGG) was left undigested (lane 4), incubated with *Pvu* I that generated two predicted fragments (lane 5), or incubated with *Ava* I that failed to cut as expected (lane 6). The digestion patterns confirmed the existence of two separate cDNA clones.
Figure 2.2 The deduced amino acid sequences of DmNopp140-RGG and DmNopp140. A CT22845, which encodes DmNopp140-RGG, consists of three exons. The first, second, and third exons encode amino acids 1-34, 35-551, and 552-688, respectively. The serine-rich acidic regions are highlighted in light gray. MPF phosphorylation motifs are in bold. The RGG domain spans residues 612-669. Two additional RGG motifs reside at 562-567.

B CT22833, which encodes DmNopp140, consists of four exons. The first two exons are identical to those in CT22845. Thus, the two proteins are identical up to residue 551 (dark gray box), after which their sequences diverge. The third exon in CT22833 encodes residues 552-604, and the fourth exon encodes residues 605-654. The carboxy terminus of DmNopp140 (the last 110 residues) is 58% identical to the carboxy terminus of rat Nopp140 (the last 106 residues). A highly conserved serine residue within this terminus (residue 610 in the black box) is a putative substrate site for cAMP dependent protein kinase.
<table>
<thead>
<tr>
<th>Rat</th>
<th>Human</th>
<th>Frog</th>
<th>Fly</th>
<th>Yeast</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TKAL</td>
<td>VKPV</td>
<td>E</td>
<td>S</td>
</tr>
<tr>
<td>2</td>
<td>RLQV</td>
<td>PKPK</td>
<td>D</td>
<td>S</td>
</tr>
<tr>
<td>3</td>
<td>NOAE</td>
<td>ognev</td>
<td>T</td>
<td>S</td>
</tr>
<tr>
<td>4</td>
<td>DSSSEEDEPPPKRACPAOGQKPVVKAVQPKAAKSSSSE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>GGEEAA</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>TSQK</td>
<td>PRP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>TAEW</td>
<td>TK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>EDAV</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>ALDK</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>YKAP</td>
<td>Z</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ACPS</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>FIPT</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>SEKK</td>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>EKEE</td>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>DAKP</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>AYEE</td>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>ESDS</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>TSSE</td>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>KSMQ</td>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>ALAV</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>NQPA</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>PAKA</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>EAPK</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>AEPK</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>KAPK</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>KAPK</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>KSAT</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>SSK</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>NMP</td>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>TPK</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>SSSS</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>SSSS</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>SSSS</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>SSSS</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>SSSS</td>
<td>S</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2.3** Amino acid sequence homology of rat Noppl40, human Noppl40, frog (*X. laevis*) Noppl40, fly Noppl40 (DmNopp140), and yeast Noppl40 (SRP40).
5.5 kbps downstream of the 3' end of *D. melanogaster Nopp140* (Flybase, 1999). This P-element could be used in the future to create a chromosomal deficiency that could potentially eliminate the *D. melanogaster Nopp140* gene.

CG7421 was predicted by the computer programs Genie and Genescan. Genie and Genescan find genes from sequence data by applying discrete Fourier transform algorithms in order to detect three base correlations (Tiwari et al., 1997). These three base correlations correspond to the coding sequence of a gene. Genie and Genescan search for this periodicity at specific frequencies and detect a candidate coding sequence. The programs then scan the sequence to locate the proper start and stop codons and verify that the coding region is in fact a gene.

Interestingly, these programs predicts that CG7421 encodes two transcripts, conceptual transcript no. 22833 (CT22833) and CT22845. CT22845 encodes DmNopp140-RGG, and it results from the expression of three exons from the pre-mRNA. These exons include nucleotides 88-186, 389-1942, and 2580-2993 of the pre-mRNA. These exons correspond to amino acid residues 1-33, 34-551, and 552-688, respectively. The deduced amino acid sequence of DmNopp140-RGG is shown in Fig. 2.2A. CT22833 on the other hand, encodes the true *D. melanogaster Nopp140* protein, that we call DmNopp140. Unlike DmNopp140-RGG, DmNopp140 is encoded by four exons corresponding to nucleotides 88-186, 389-1942, 2184-2342, and 3096-3248 of the pre-mRNA. These exons encode amino acid residues 1-33, 34-551, 552-604, and 605-654, respectively. The deduced amino acid sequence of DmNopp140 is shown in Fig. 2.2B. Therefore, CT22845 (DmNopp140-RGG) and CT22833 (DmNopp140) are identical in their first two exons (nucleotides 88-186 and 389-1942 of the pre-mRNA). Due to alternative splicing of the mRNA, however, these two transcripts encode carboxy
termini that are completely different. The carboxy terminal tail of DmNoppl40 is highly conserved among previously identified Nopp140 proteins (Meier 1996, see Fig. 2.3, 2.4). The last 110 amino acid residues of DmNoppl40 are 58% identical to the last 106 amino acid residues of rat Nopp140. According to Issac et al. (1999) the carboxy terminal region of rat Nopp140 can be subdivided into two distinct domains, Ca (residues 579-655) and Cb (residues 656-704). Cb is present in Nopp140 proteins of all species, while Ca is present only in Nopp140 proteins of metazoans. Consistent with this observation, DmNoppl40 contains both the Cb and Ca subdomains. Interestingly, the breakpoint between the third and fourth exons that encode DmNoppl40 coincide well with Ca (residues 552-604) and Cb (residues 605-654) of the DmNoppl40 protein. DmNoppl40’s Ca has 50% identity to the Ca sub-domain in both human and rat Nopp140. DmNoppl40’s Cb sub-domain has 78% and 76% identity to the Cb subdomains of human and rat Nopp140s, respectively.

The third exon of CT22845 (DmNoppl40-RGG), nucleotides 3096-3248 of the pre-mRNA, encodes an RGG domain that is defined as arginine-glycine-glycine (RGG) repeats interspersed with aromatic amino acids (Ginisty et al., 1999). The overall length of the RGG domain in DmNoppl40-RGG is 58 amino acid residues. It consists of 44 glycine, 10 arginine, and 4 phenylalanine residues. All ten arginine residues are followed by at least two glycine residues.

DmNoppl40-RGG and DmNoppl40 contain many putative kinase phosphorylation sites. The acidic repeats in the central domain of both D. melanogaster proteins consist of primarily serine, glutamic acid, and aspartic acid residues (see Fig. 2.1A and B). The serine residues present in the acidic repeats of DmNoppl40 and DmNoppl40-RGG constitute many putative casein kinase II (CKII) phosphorylation...
Figure 2.4 Linear representations of Nopp140 homologues. A Acidic (black boxes) and basic (white boxes) regions alternate within the central domains of all Nopp140 homologues. Numbers in the carboxy regions of all proteins represent the percentage of amino acid residues that are identical to the carboxy region (last 106 residues) of the prototypical rat Nopp140. Residues compared for each homologue were: Human (last 106 residues), X. laevis (last 105 residues), D. melanogaster (last 110 residues), yeast (last 100 residues). B DmNopp140-RGG is not a true Nopp140 homologue. DmNopp140-RGG contains a carboxy domain rich in arginine-glycine-glycine (RGG) repeats instead of the prototypical evolutionarily conserved carboxy domain.
sites. DmNopp140 and DmNopp140-RGG also share eight putative Mitosis Promoting Factor (MPF) phosphorylation sites (see Fig. 2.2A and B). These sites are distributed throughout the acidic and basic repeat domains in both proteins. DmNopp140 contains a putative site for protein kinase A (PKA) phosphorylation (serine 610). Due to mRNA splicing, however, DmNopp140-RGG does not contain this PKA phosphorylation site.

Discussion

Several non-ribosomal nucleolar proteins were identified in a screen to identify proteins that bind with high affinity to a peptide probe containing the NLS of SV40 large T antigen in vitro (Osborne and Silver, 1993). All shared acidic and serine residues, including casein kinase II phosphorylation sites interspersed with stretches of basic residues. This class of proteins includes yeast NSR1 and NPI46, and mammalian B23, nucleolin, and Nopp140. Osborne and Silver (1993) propose that the acidic domains mediate interactions with NLS containing proteins through electrostatic interactions. In addition, these interactions may depend upon phosphorylation states (Xue and Melese, 1994) (see Chapter 3 for a more detailed discussion on phosphorylation).

The overall domain structure of the D. melanogaster variants and rat Nopp140 are similar in that both consist of central alternating acidic and basic domains, but only DmNopp140 contains a conserved carboxy terminus. One difference between the D. melanogaster and rat proteins is that the D. melanogaster proteins contain more acidic domains than that of rat Nopp140 (15 as compared to 12, see Fig. 2.4A). The X. laevis homologue of Nopp140, called xNopp180, contains 18 acidic regions. It is interesting to note, however, that although the D. melanogaster variants contain more acidic domains than that of rat Nopp140, the overall length of their central acidic/basic domain is shorter. The significance of these differences has yet to be determined.
The serine residues present in the acidic repeats of DmNopp140 and DmNopp140-RGG constitute many putative casein kinase II (CKII) phosphorylation sites (see Chapter 3). A CKII consensus phosphorylation site consists of a serine or threonine acceptor site and an acidic amino acid three residues away on its carboxy-terminal side. Every acidic amino acid residue near the target serine improves the acceptor site (Marin et al., 1986; Kuenzel et al., 1987). Phosphorylation of the target serine within an acidic repeat causes the preceding serine to become a better substrate for CKII. Due to the high number of putative CKII phosphorylation sites present in both *D. melanogaster* proteins, DmNopp140 and DmNopp140-RGG should serve as excellent substrates for CKII (see Chapter 3).

DmNopp140 and DmNopp140-RGG share eight putative Mitosis Promoting Factor (MPF) phosphorylation sites (see Fig. 2.2A and B). These sites are distributed throughout several of the basic repeat domains in both proteins. The human homologue of Nopp140 (p130) was shown to be hyperphosphorylated during mitosis, probably by MPF (Pai et al., 1995). MPF reportedly modulates the activity of other nucleolar proteins such as nucleolin (Belenguer et al., 1990; Peter et al., 1990) and B23 (Peter et al., 1990). Due to the putative MPF sites present in DmNopp140 and DmNopp140-RGG, both proteins should serve as substrates for this enzyme (see Chapter 3).

DmNopp140 contains a putative site for protein kinase A (PKA) phosphorylation (serine 610). Serine 610 is present in the carboxy terminal domain of DmNopp140, but is absent in DmNopp140-RGG due to the mutually exclusive splicing of the mRNA. Therefore, DmNopp140-RGG does not share this putative PKA phosphorylation site. This suggests that DmNopp140 may be regulated by signal transduction cascades that do not regulate DmNopp140-RGG.
DmNopp140-RGG's RGG domain is similar to RGG domains present in other proteins such as mammalian GAR1p, fibrillarin, and nucleolin. The RGG domain of CHO nucleolin is 52 residues in length consisting of 37 glycine, 10 arginine, and 5 phenylalanine residues. In Chinese hamster ovary (CHO) nucleolin, nine of the arginine residues are followed by at least two glycine residues. The RGG domain of DmNopp140-RGG is 58 residues long consisting of 44 glycine, 10 arginine, and 4 phenylalanine residues. In DmNopp140-RGG, ten of the arginine residues are followed by at least two glycine residues. There is another arginine and glycine rich domain towards the amino-terminal end of DmNopp140-RGG that contains only two RGG repeats (residues 562-567). This small RGG domain may function as large RGG domains or it may function in concert with the larger RGG domain present in DmNopp140-RGG. Interestingly, in most proteins that contain an RGG domain, the domain is located near its carboxy terminal end. GAR1p however, contains RGG domains both at its amino and carboxy terminal end. The RGG domain of fibrillarin is located in its amino terminus. The significance of placement of the RGG domain in a protein and the function of the RGG domain itself have not yet been determined (see Chapter 1 for a more detailed discussion of RGG domains). As will be discussed in later chapters, the differences in the carboxy ends of the two D. melanogaster proteins confer noticeably different localization patterns in HeLa cells.

In summary, the D. melanogaster Nopp140 gene encodes two proteins, DmNopp140 and DmNopp140-RGG. DmNopp140 is the bona fide homologue of vertebrate Nopp140 in that it shares an evolutionarily conserved carboxy terminus. DmNopp140-RGG contains a carboxy terminus that contains an extensive RGG domain in place of the evolutionarily conserved carboxy terminus.
Chapter 3
Expression and Phosphorylation of DmNopp140-RGG

Introduction

Phosphorylation is important for many cellular processes such as cell cycle regulation, signal transduction, transcription, and translation. Phosphorylation is vital for fundamental processes of the nucleolus as well. In fact, rDNA transcription by RNA polymerase I is greatly influenced by the phosphorylation of specific nucleolar proteins. Transcription of rDNA by RNA Pol I oscillates during the cell cycle, reaching a maximum in S phase and G2, diminishing during mitosis, and then recovering during G1 progression. Two specific nucleolar proteins, upstream-binding factor (UBF) and the transcription initiation factor (TIF) –IB/SLI are both required for Pol I transcription initiation. Both are modulated by phosphorylation (Klein and Grummt, 1999). UBF is phosphorylated by casein kinase II (CKII) at serine residues within its carboxy terminal acidic domain. CKII phosphorylation of UBF is absolutely required for transcriptional activation (Voit et al., 1995b). Yet, examination of UBF mutants in which the serine residues within its C-terminal tail were altered by site specific mutagenesis demonstrated that CKII – mediated phosphorylation is necessary, but not sufficient for transcriptional activation (Voit et al., 1995a). These results suggest that other factors are needed for activation of RNA Pol I transcription.

Phosphorylation of UBF can also serve to negatively regulate RNA Pol I transcription. For example, MPF phosphorylation of SLI and UBF represses Pol I transcription during mitosis (Klein and Grummt, 1999). Repression of RNA Pol I transcription during cell cycle arrest may be due to a phosphorylation and dephosphorylation cycle.
When 3T6 cells become confluent, they withdraw from the cell cycle. As cell confluency increases, the hypophosphorylated form of the tumor suppressor protein, Rb, increases and accumulates within nucleoli. Co-immunoprecipitation experiments demonstrate that the interaction of hypophosphorylated Rb and UBF increases. This increased interaction between hypophosphorylated Rb and UBF correlates with the reduction of rDNA transcription (Hannan et al., 2000).

Other nucleolar proteins besides those involved in basal RNA Pol I transcription are modulated by phosphorylation. Nucleolin, a protein that interacts with pre-rRNA to perhaps facilitate early cleavage events and ribosome assembly (see Chapter 1), contains several amino-terminal acidic domains that are rich in serine residues. This region is highly phosphorylated (Bourbon et al., 1983; Rao et al., 1982). Nucleolin is a substrate for CKII (Caizergues – Ferrer et al., 1987) and p34\(^{\text{cdc2}}\) (MPF) (Belenguer et al., 1990; Peter et al., 1990), and its phosphorylation is highly regulated by the cell cycle. B23, another phosphoprotein of the nucleolus, has ribonuclease and chaperone activities (Szebeni and Olson, 1999). B23 is a substrate for CKII (Chang et al., 1990) and MPF (Peter et al., 1990). Interestingly, B23 is also a substrate of CDK2/cyclin E during centrosome duplication (Okuda et al., 2000). A fraction of B23 associates with unduplicated centrosomes until it is phosphorylated by CDK2/cyclin E at which time it dissociates from the centrosome.

Nopp140 was first discovered as an NLS binding protein that localizes to the nucleolus of Buffalo rat liver cells (Meier and Blobel, 1990). Nopp140 shuttles between the nucleolus and cytoplasm, and its NLS-containing peptide-binding activity is dependent upon its phosphorylation (Meier and Blobel, 1992). Nopp140 is unique due to its extensive phosphorylation. Phosphatase treatment of \textit{in vitro} translated rat Nopp140
shifts its apparent molecular weight from 140 kDa to 100 kDa on SDS-PAGE gels (Meier and Blobel, 1992). Conversely, treatment of recombinant rat Nopp140 with CKII shifts its apparent molecular weight on SDS-PAGE gels from 100 kDa to 140 kDa (Meier, 1996). Examination of the primary protein sequence of Nopp140 provides a clue as to why CKII phosphorylates Nopp140 to such a high degree (Fig. 2.1). Rat Nopp140 contains 49 putative CKII phosphorylation sites. Upon their phosphorylation, an additional 33 sites are formed. A CKII consensus phosphorylation site consists of a serine or threonine acceptor site and an acidic amino acid three residues away on its carboxy-terminal side. Every acidic amino acid residue near the target serine improves the acceptor site (Marin et al, 1986; Kuenzel et al, 1987). Phosphorylation of the target serine within a Nopp140 acidic repeat causes the preceding serine to become a better substrate for CKII. This process could occur at all of the acidic repeats, leading to massive phosphorylation of Nopp140. In addition, CKII phosphorylates SRP40, the yeast homologue of Nopp140 (Meier, 1996). CKII phosphorylation of Nopp140 homologues, therefore, appears to be highly conserved.

CKII has been observed in the nucleus and in the nucleolus in a growth dependent manner (Belenguer et al., 1989). Other non-ribosomal nucleolar proteins such as nucleolin, B23, and the Nopp140 associated protein of 57 kDa (NAP57) are substrates of CKII (Xue and Melese, 1994). All of these proteins have common stretches of alternating acidic and basic domains in which the serines within the acidic domains serve as excellent substrates for CKII phosphorylation. Like Nopp140, these proteins all bind NLS containing peptides, and this NLS-binding activity is dependent upon CKII phosphorylation (Xue and Melese, 1994).
CKII is a heterotetramer where α and ω are catalytic subunits, and β is the regulatory subunit. The predominant forms are α₂β₂, αωβ₂, or αβ₂. CKII phosphorylation has been implicated in nucleic acid synthesis, signal transduction, transcription, and translation (Allende and Allende, 1995). CKII may play an important role in cell growth and division. For example, CKII is required for progression of the cell cycle (Pepperkok et al., 1994; Hanna et al., 1995). CKII also associates with the growth-related proteins p53 (Filhol et al., 1992) and c-Raf (Janosch et al., 1996).

CKII not only phosphorylates nucleolin, but it also forms a specific interaction with nucleolin in vitro, and perhaps in vivo (Li et al., 1996). A specific interaction between CKII and Nopp140 has also been observed (Li et al., 1997). This interaction with Nopp140, as shown by immunoprecipitation, is mediated through the β regulatory subunits of CKII. The interaction with nucleolin is mediated through the α subunit of CKII. Specific interactions between nucleolar proteins and the various subunits of CKII may represent another level of regulation for these proteins besides phosphorylation.

A portion of human Nopp140 (hNopp140) fluctuates in synchrony with the cell cycle (Pai et al., 1995). Interphase hNopp140 migrates on SDS-PAGE gels at 130 kDa but also at 95 kDa. The conversion of the 95 kDa species to the 130 kDa species is due to CKII phosphorylation. Interestingly, another hNopp140 species that migrates slower than 130 kDa on SDS-PAGE gels exists in mitotic extracts of CEM (human T leukemic) cells and HeLa cells (Pai et al., 1995). This hyperphosphorylation, due to MPF, increases the solubility of Nopp140 and disperses it throughout the entire mitotic cytoplasm.

MPF is a heterodimer composed of a mitotic cyclin (cyclin B) and a cyclin dependent protein kinase (p34cdk2 kinase or CDK1). The kinase activity of MPF phosphorylates multiple protein substrates as its activity rises at the onset of mitosis in
order to push the cell into mitosis. Some of the protein targets of MPF are the nuclear lamins (Heald and McKeon, 1990), condensin complexes (Biggins and Murray, 1998), and myosin light chains (Nigg et al., 1996). MPF phosphorylation of the nuclear lamins leads to nuclear envelope breakdown. Condensin complexes, when phosphorylated by MPF, lead to chromosome condensation. MPF-catalyzed phosphorylation of the myosin light chain prevents cytokinesis. Cytokinesis does not occur until MPF activity falls at the completion of anaphase. Thus, MPF plays an important role in the regulation of events that are involved with mitosis.

Although most targets of MPF are unknown, there are several known target proteins present in the nucleolus. MPF phosphorylation of nucleolin occurs on threonine residues in the basic repeats of its N-terminal domain. MPF, however, may not phosphorylate all potential sites in vivo as efficiently as it does in vitro (Belenguer et al., 1989; Peter et al., 1990). Human Nopp140 is a target of MPF activity (Pai et al., 1995). MPF may hyperphosphorylate hNopp140 during mitosis in order to increase its solubility as in the case of the nuclear lamins (Gerace and Blobel, 1980; Peter et al., 1990).

Chapter 3 describes the subcloning and expression of DmNopp140-RGG, as well as its in vitro phosphorylation by CKII and MPF.

Materials and Methods

Sub-cloning of DmNopp140-RGG into pET30

The full-length cDNA encoding DmNopp140-RGG was removed from LD10913 (see Chapter 2) using a Bsp HI site located at the translation start site and a Bam HI site located within the 3' UTR (untranslated region). The cDNA was ligated into pET30 (Novagen) at the Nco I site, which is compatible with Bsp HI, and at the Bam HI site such that the DmNopp140-RGG coding sequence was positioned downstream and in frame.
with the sequence of pET30 that encodes the six-histidine tag (His-tag). This recombinant plasmid, pET30/DmNopp140-RGG, was transformed into the *E. coli* strain JM109 DE3 pLysS.

**Expression of DmNopp140-RGG**

For over-expression of DmNopp140-RGG, a 5 ml Luria-Bertani (LB) culture supplemented with 30 µg/ml kanamycin was inoculated with a single colony of pET30/DmNopp140-RGG in JM109 DE3 pLysS from a streaked plate. The next day, 5 ml of this overnight culture was transferred to 1 L of LB broth supplemented with 30 µg/ml kanamycin. The culture was grown to an OD<sub>600</sub> of 0.6, at which point isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. The culture flask was then transferred to a shaking water bath at a temperature of 28°C and allowed to incubate overnight at a setting of 500 RPM. The cells were harvested by centrifugation at approximately 6000 x g, and the pellets were resuspended in 30 ml of His binding buffer (5 mM imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9) supplemented with 0.1 mM PMSF, 1.0 µg/ml pepstatin, and 1.0 µg/ml leupeptin. Cells were lysed by sonication for 2 minutes on ice. The cell lysate was cleared by centrifugation at 25000 x g in a SW50.1 rotor for 20 minutes at 4°C. The supernatant was passed through a 0.4 µm syringe filter directly onto a His-tag column using Novagen’s recommendations. The columns were washed with 10 bed volumes of 1X binding buffer and 6 bed volumes of 1X wash buffer (60 mM imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9). The protein was eluted with 6 bed volumes of 1X elution buffer (1 M imidazole; 0.5 M NaCl; 20 mM Tris-HCl, pH 7.9). All buffers were supplemented with 0.1 mM PMSF, 1.0 µg/ml leupeptin, and 1.0 µg/ml pepstatin. Eluate was collected and dialyzed overnight at 4°C in 75 mM KCl, 10 mM Tris-HCl, pH 7.2, 1 mM EDTA.
supplemented with 0.1 mM PMSF. The protein was concentrated by placing the dialysis bag (molecular weight cut-off of 13000 to 14000) into solid polyvinylpyrrolidone (PVP, average molecular weight= 360,000) at 4° C. This resulted in rapid dehydration without adversely changing ionic strength.

**Phosphorylation of DmNopp140-RGG**

The enriched protein was phosphorylated *in vitro* using gamma-labeled [32P] - ATP at 800 Ci/mmol (ICN Pharmaceuticals, Costa Mesa, CA) and either casein kinase II (CKII) or p34<sup>Cdc2/cyclinB</sup> protein kinase (MPF). Both enzymes were purchased from New England Biolabs (Beverly, MA). Protein was incubated with gamma-labeled [32P] - ATP and either CKII or MPF in a 1X reaction buffer (supplied by manufacturer) for 1 hour at 37° C. After addition of 1X SDS sample buffer, the samples were boiled for 2 minutes, and then applied to an 8% PAGE gel containing SDS. The gel was run at 100 V for 1.5 hours to resolve the protein. The gel was stained with Coomassie Blue, and then destained overnight in Coomassie destain solution. The destained gel was dried onto Whatman 3 MM paper using a gel dryer. After the gel was completely dry, it was applied to high-speed X-ray film (Kodak) for 1 hr.

**Results**

The deduced peptide sequence for DmNopp140-RGG indicated 66 putative CKII phosphorylation sites (Fig. 2.1 A). Upon their phosphorylation an additional 18 sites are formed. To demonstrate that DmNopp140 is an *in vitro* substrate for CKII, DmNopp140-RGG was expressed as an amino-terminal His tagged fusion protein in *Escherichia coli*. The fusion protein was purified over a nickel affinity column. The protein was labeled *in vitro* by CKII in the presence of gamma [32P] - ATP (Fig. 3.1).
Figure 3.1 *In vitro* CKII and MPF phosphorylation of DmNopp140-RGG. The panel on the left is a Coomassie stained gel showing His-tagged DmNopp140-RGG that was expressed in *E. coli* and purified using nickel affinity columns. Proteolysis was unavoidable. The panel on the right is the corresponding autoradiogram. Lanes A and A’ Purified DmNopp140-RGG incubated with MPF and [*32P*-gamma-labeled ATP. Lanes B and B’ Purified DmNopp140-RGG incubated with CKII and [*32P*-gamma-labeled ATP. Lanes C and C’ Purified DmNopp140-RGG incubated with [*32P*-gamma-labeled ATP, but with no enzyme. The same amount of protein was used in each assay. CKII phosphorylation greatly exceeds MPF phosphorylation, probably due to the greater number of CKII sites versus MPF sites. Excess phosphorylation by CKII resulted in a detectable shift in molecular weight (see arrow).
The deduced peptide sequence of DmNopp140-RGG indicated eight putative MPF phosphorylation sites within the basic repeats (bold SPKK or TPAK sites in Fig. 2.2A). These sites are identical in DmNopp140 (Fig. 2.2B). To demonstrate that DmNopp140-RGG was an \textit{in vitro} substrate for MPF, DmNopp140-RGG was expressed as an amino-terminal His tagged fusion protein in \textit{E coli}. The fusion protein was purified over a nickel affinity column. The protein was labeled \textit{in vitro} by MPF in the presence of gamma $^{32}$P – ATP (Fig. 3.1).

\textbf{Discussion}

CKII phosphorylated affinity-purified DmNopp140-RGG to a much greater degree than did MPF as demonstrated by the intensity of the signals captured on the autoradiogram (compare A’ to B’). In addition, phosphorylation by CKII caused a major shift in the position of the highest molecular weight band of DmNopp140-RGG (full-length protein). Phosphorylation by MPF did not cause this major shift (as compared to C and C’). This is most likely due to the higher number of putative CKII phosphorylation sites as compared to MPF phosphorylation sites (see Fig. 2.2A and B).

To ensure that the autoradiograph signals observed were specific for DmNopp140-RGG, CKII, and MPF, labeling reactions were performed with no enzyme. DmNopp140-RGG was not labeled without the presence of either CKII or MPF (Fig. 3.1 C and C’). In addition, labeling reactions were performed with no protein present to determine if the signal was the enzyme labeling itself. It was determined that CKII could label itself, but that the signal obtained in no way resembled the signal obtained from \textit{in vitro} labeling reactions containing DmNopp140-RGG and CKII or MPF (not shown). There was no labeling detected in \textit{in vitro} labeling reactions that contained MPF without DmNopp140-RGG (not shown).
There was a major shift in the position of DmNopp140-RGG when labeled with CKII. DmNopp140-RGG labeled by CKII migrated slower on polyacrylamide gels as compared to DmNopp140-RGG that was not labeled (compare Fig. 3.1B and C). Meier and Blobel (1992) observed a reverse shift for \textit{in vitro} translated rat Nopp140 when exposed to alkaline phosphatase. In this case, \textit{in vitro} translated rat Nopp140 shifted from a phosphorylated state (M, 140 kDa) to an unphosphorylated state (M, 100 kDa). Conversely, further incubation of the translation products in reticulate lysate after translation arrest by dilution with buffer resulted in conversion of the 100 kDa protein to the 140 kDa species of Nopp140.

Interestingly, the phosphorylation of rat Nopp140 was "an all or none" phenomenon and affected the ability of the protein to bind NLS-containing peptides (Meier and Blobel, 1992). The shift observed with \textit{E. coli} expressed DmNopp140-RGG was not as dramatic. DmNopp140-RGG shifted only from \~{}100 kDa to \~{}120 kDa (Fig. 3.1 B).

It was expected that the shift of DmNopp140-RGG would be greater than that of rat Nopp140 based upon the fact that DmNopp140-RGG contains 64 putative CKII phosphorylation sites as compared to 49 in rat Nopp140. Not all of the theoretical CKII phosphorylation sites may have been utilized in the \textit{in vitro} phosphorylation reactions. Alternatively, proteolysis may have prevented the accumulation of full-length phosphorylated DmNopp140-RGG.

The severe proteolysis of \textit{E. coli} expressed DmNopp140-RGG, as visualized on polyacrylamide gels (not shown), was unavoidable. This could be due to post-translational modification steps needed to stabilize the protein that were absent in \textit{E. coli}. Protocol steps have been taken to remedy this problem. For example, expression of a
GFP-tagged version of DmNopp140-RGG in an *in vitro* transcription/translation coupled system was successful as no proteolysis was observed (not shown).

An interesting comparison can be drawn between nucleolin and Nopp140. Both proteins are phosphorylated by CKII, and subunits of CKII interact directly with both proteins (Caizergues-Ferrer et al., 1987; Li et al., 1997). Both nucleolin and Nopp140 can also be phosphorylated by MPF (Belenguer et al., 1989; Peter et al., 1990; Pai et al., 1995; this report). Finally, nucleolin is a potential substrate for protein kinase C (Zhou et al., 1997), while Nopp140 is a potential substrate for protein kinase A (Meier, 1996). This suggests that nucleolin and Nopp140 are regulated in a similar manner. Interestingly, although all Nopp140 homologues, including DmNopp140, contain a conserved serine that is a putative protein kinase A site, this site is missing in DmNopp140-RGG. This suggests a slightly different regulatory mechanism for the true *Drosophila* Nopp140 homologue over the RGG variant. This may indicate different functions for the two *Drosophila* Nopp140 splice variants.
Chapter 4

Localization Patterns of DmNopp140-RGG and DmNopp140 Within Nucleoli and CBs

Introduction

The research performed in Chapter 4 establishes the localization patterns of DmNopp140-RGG and DmNopp140 in *Drosophila melanogaster* Schneider II cells, HeLa cells, and in *Xenopus laevis* oocytes. This Chapter will also examine localization patterns of the endogenous nucleolar proteins fibrillarin and nucleolin with respect to both *D. melanogaster* Nopp140 variants in the same cell lines. Localizations between *X. laevis* coilin and both *D. melanogaster* Nopp140 variants within CBs will also be addressed.

Over the years, many interactions have been observed between nucleoli and CBs. CBs were originally called “nucleolar accessory bodies” by Ramon y Cajal who discovered these nuclear bodies in 1903 (Cajal, 1903) due to their close association with the periphery of nucleoli in neurons. In fact, in certain cases, including human breast carcinoma cells, brown adipocytes, and hepatocytes of hibernating dormice, CBs have been detected inside of the nucleolus (Ochs et al., 1994; Malatesta et al., 1994). This is not to say that CBs are always associated with nucleoli, since CBs in most cells occur in the nucleoplasm, separate from nucleoli. A recent report demonstrates that CBs are highly mobile, and can join and separate from each other, and move to and from the nucleolus (Platani et al., 2000).

CBs contain a number of nucleolar proteins and RNA species. The nucleolar proteins fibrillarin, Nopp140, Nap57, and the ribosomal protein S6 along with the snoRNAs U3, U8, and U14 all have been discovered in CBs (Bohmann et al., 1995a;
Narayanan et al., 1999; Gall, 2000). CBs, at least in *X. laevis* oocytes, are known to contain all three RNA polymerases, TFIIA, and RAP74, a subunit of TFIIF. Gall et al. (1999) have proposed a model in which CBs function as a staging area for the production of RNA polymerase transcription and processing complexes before their transport to active sites of transcription (see Chapter 1).

CBs contain proteins that are intimately associated with snoRNP's. These include fibrillarin, Nopp140, and NAP57 (Yang et al., 2000). The interaction between fibrillarin and box C/D snoRNPs is well established (Tyc and Steitz, 1989). These results suggest a role for CBs in snoRNP biogenesis or function.

A functional link between the nucleolus and CBs exists through the CB marker protein, p80 coilin. Antibodies directed against p80 coilin can be used to unambiguously identify CBs in many different species including plants (Beven et al., 1995). P80 coilin shows no homology to any other protein except for the *X. laevis* derived SPH-1, which is now accepted as a homologue of p80 coilin (Tuma et al., 1993). Certain p80 coilin mutants, when exogenously expressed in HeLa cells, caused segregation of antigens from the fibrillar components of the nucleolus (Bohmann et al., 1995b). Further, exposure of HeLa cells to low levels of okadaic acid (a serine/threonine protein phosphatase inhibitor), or expression of a p80 coilin mutant (S202D) caused relocation of p80 coilin and splicing snRNPs to the nucleolus. Taken together, these experiments point to a functional relationship between the nucleolus and CBs.

The only nucleolar protein known to physically interact with p80 coilin is Nopp140. Not only did expression of a partial Nopp140 construct lead to dispersal of CBs, but p80 coilin was identified in a yeast two-hybrid screen to interact with Nopp140 (Isaac et al., 1998). The partial Nopp140 construct used in the screen (HA-NoppC)
consisted of only the carboxyl terminal domain corresponding to Ca and Cb (see Chapter 2). Exogenously expressed HA-NoppC localized to the nucleoplasm without accumulation in the nucleolus. Interestingly, in HA-NoppC transfected cells, endogenous Nopp140 as well as NAP57 and fibrillarin were chased out of the nucleolus. Nucleolar localization of nucleolin, UBF, B23, and RNA Pol I were unaffected; they remained within residual nucleoli. Another effect that HA-NoppC expression had was to disperse CBs. In HA-NoppC transfected cells, both p80 coilin and Sm antigens diffused from CBs into the nucleoplasm.

Yeast two-hybrid analysis demonstrated that the amino terminus of p80 coilin also interacts with Nopp140. This amino terminus contains the sequence required for its targeting to CBs in amphibian oocytes (Wu et al., 1994). The entire sequence of Nopp140 appears to be required for the interaction with p80 coilin, since any truncation of Nopp140 resulted in a complete loss of the interaction.

Experiments performed to study the targeting of rat Nopp140 to the nucleolus and CBs show that exogenously expressed Nopp140 accumulates in nucleoli 6-12 hours after transfection (Isaac et al., 1998). Only after a lag phase of an additional 12 hours is Nopp140 observed in CBs. This suggests that newly synthesized Nopp140 first targets to nucleoli, and then to CBs, indicating that the path to CBs is through the nucleolus. Deletion analysis of Nopp140 demonstrates that the conserved central repeat domain localizes initially to nucleoli, then it is excluded from nucleoli after 48 hours. COS-1 cells expressing only the amino terminus or carboxy terminus of Nopp140 show that neither accumulate in the nucleolus or CBs. These results suggest that the central repeat domain is required for the localization of Nopp140 to nucleoli and CBs. Expression of the central repeat domain alone, however, causes the formation of ring-like structures.
named Nopp-R induced rings (R-rings). These R-rings attract or redirect fibrillarin, NAP57, UBF and RNA Pol I, while not affecting the nucleolar localization of nucleolin and B23. Incidentally, the nucleoli within cells containing R-rings remained intact on a light microscopic level.

Based upon amino acid sequence similarities between DmNopp140 and rat Nopp140, DmNopp140 should localize to nucleoli and CBs. DmNopp140-RGG should localize to nucleoli based upon its similarities with Rat Nopp140 and nucleolin. The RGG domain of nucleolin was shown to be necessary but not sufficient for its nucleolar localization (Heine et al., 1993). DmNopp140-RGG shares homology with rat Nopp140, which localizes to CBs. DmNopp140-RGG also shares homology with nucleolin, which does not localize to CBs. This chapter presents the localizations of DmNopp140-RGG and DmNopp140 in D. melanogaster Schneider II cells, HeLa cells, and X. laevis oocytes as they compare with exogenous coilin and endogenous fibrillarin and nucleolin.

**Materials and Methods**

**Construction of Expression Vectors**

**DmNopp140-RGG**

In order to follow in vivo localizations of DmNopp140, DmNopp140-RGG, and X. laevis coilin without antibodies against these proteins, vectors were constructed in which these proteins could be expressed as either green fluorescent protein (GFP) or red fluorescent protein (RFP) fusions (see Appendix B for maps of all constructs used for localizations). The two expression vectors used were pEGFP-C3 (encoding enhanced green fluorescent protein from the jellyfish *Aequoria sp.* ) and pDsRed1-Cl (encoding red fluorescent protein from an Indopacific sea anenome relative *Discoma sp.* ) (Clontech, LaJolla, CA).
The full-length cDNA encoding DmNoppl40-RGG was excised from LD10913 (see Chapter 2 for description) at the Bsp HI and Bam HI restriction sites and inserted into the multiple cloning site (MCS) of pET30 at Nco I (complimentary to Bsp HI) and Bam HI. A cDNA insert corresponding to the full-length DmNoppl40-RGG cDNA and the enterokinase site from pET30 was then cut from pET30/DmNoppl40-RGG at a Bgl II site upstream of the translation start site and an Eco R1 site downstream of the translation stop site. This insert was ligated into the pBluescript KS+ vector. This plasmid was used to synthesize capped mRNA that encodes GFP-tagged DmNoppl40-RGG.

To place DNA encoding GFP-tagged DmNoppl40-RGG behind a strong mammalian CMV promoter, DNA encoding DmNoppl40-RGG behind the EGFP tag was cut from this construct with Sac I upstream and Eco R1 downstream of the insert. This insert was ligated into the expression vector pBKCMV at the same restriction sites. This construct also has T7 and T3 promoters to synthesize antisense and sense mRNAs (respectively) for X. laevis oocyte microinjections. Unfortunately, in vitro transcription reactions using pBKCMV/EGFP-DmNoppl40-RGG as a template were failures. On the other hand, this construct could be used for transfection of mammalian cell lines (HeLa, CHO). Since effort to synthesize sense mRNAs encoding EGFP-DmNoppl40-RGG were unsuccessful from the T3 promoter, efforts were made to use the T7 promoter. DNA encoding EGFP-DmNoppl40-RGG was cut from pBKCMV/EGFP-DmNoppl40-RGG at Sac I upstream and Eco R1 downstream of the DNA. This insert was ligated into pBluescript KS+ at the same sites. This construct was named pBluescript KS+/EGFP-DmNoppl40-RGG.
In order to construct a DsRed-tagged version of DmNopp140-RGG, DNA encoding DmNopp140-RGG was cut from the pET30 construct at a Bgl II site upstream of the translation start site and at an Eco RI site downstream of the translation stop site. This insert was ligated into pDsRed1-C1 at the same restriction sites, and the construct was named pDsRed1-C1/DmNopp140-RGG (see Appendix).

DmNopp140

The full-length cDNA encoding DmNopp140 was removed from SD10348 (see Chapter 2) at Bsp HI at the translation start site and Xho I at the 3’ end of the clone, and inserted into pET30 at Nco I and Xho I (Bsp HI and Nco I have compatible sticky ends). The cDNA corresponding to full-length DmNopp140 was removed from pET30 with a complete digestion by Xho I, followed by a partial digestion with Bgl II (there is a Bgl II site internal to the DmNopp140 cDNA 114 bp from the translation stop site). This DNA was inserted into pDsRed1-C1 at Bgl II and Sal I such that the coding sequence for DmNopp140 was downstream and in frame with the coding sequence for the DsRed tag. This expression vector is named pDsRed1-C1/DmNopp140 (see Appendix). The same strategy was used to clone the DmNopp140 cDNA into the pEGFP-C3 vector to construct pEGFP-C3/DmNopp140 (see Appendix).

In order to synthesize mRNAs encoding a GFP-tagged version of DmNopp140, the cDNA had to be moved into a vector with either a T3 or T7 promoter. Therefore, the DNA encoding GFP-tagged DmNopp140 was cut from the pEGFP-C3 plasmid using Nhe I (upstream of GFP tag) and Sal I (downstream of translation stop), and inserted into pBluescript KS+ at Xba I (compatible with Nhe I) and Sal I. This ligation resulted in a construct, pBluescript KS+/EGFP-DmNopp140, in which capped mRNAs encoding a GFP-tagged version of DmNopp140 could be synthesized from the T7 promoter.
**DmNopp140ΔRGG**

In order to engineer a carboxy-terminal truncation of DmNopp140-RGG (DmNopp140ΔRGG), two PCR primers were designed to amplify the DmNopp140-RGG cDNA (LD10913) but lacking the sequences encoding the RGG tail. The upstream primer was complimentary to the non-coding strand, and it contained an *Eco* RI site (bold) just upstream of the ATG start codon (underlined) for cloning purposes (5'-CGGAATTCT CATGACAGACCTGCTAAAGATAGCC-3'). The downstream primer was complimentary to the coding strand, and it contained a stop codon (underlined) that would have normally encoded amino acid residue 562 of DmNopp140-RGG (5'-AAG GATCCTTATCCGTTGTGCTTCTTAAAG TCG-3'). A *Bam* HI site (bold) was included in the downstream primer for cloning purposes. The resulting PCR product was digested with *Eco* RI and *Bam* HI, and then ligated into pEGFP-C3 at the same sites such that the coding sequence for DmNopp140ΔRGG was downstream and in frame with the coding sequence for EGFP. The resulting construct was named pEGFP-C3/DmNopp140ΔRGG (see Appendix).

**X. laevis** **Coilin**

A clone encoding a myc-tagged version of *X. laevis* coilin in pBluescript II KS+ was a gift of Zheng'an Wu in Dr. Joe Gall's lab (Carnegie Institution of Washington, Baltimore, MD). The cDNA encoding *X. laevis* coilin was removed with the restriction enzymes *Bam* HI and *Xba* I and ligated into pEGFP-C3 at the same restriction sites. This resulted in a construct, pEGFP-C3/*X. laevis* coilin, in which the coding sequence of *X. laevis* coilin was downstream and in frame with the coding sequence of the EGFP tag (see Appendix).
In order to construct a DsRed-tagged version of *X. laevis* coilin, the cDNA first had to be subcloned into pBluescript KS+ to pick up useful restriction sites. The cDNA encoding *X. laevis* coilin was removed from the myc-tagged construct with *Bam* HI and *Xba* I and inserted into pBluescript KS+ at the same sites. The *X. laevis* coilin cDNA was then cut from pBluescript KS+ with *Eco* RI at the 5’ and *Sac* II at the 3’ end of the DNA, and then inserted into pDsRedl-Cl at the same sites. This resulted in a construct, pDsRed1-C1/*X. laevis* coilin (see Appendix B), in which the coding sequence of *X. laevis* coilin was ligated downstream and in frame with the coding sequence of the DsRed tag.

**Transient Transfections of HeLa and Schneider II Cells**

All media and antibiotics were from Life Technologies (GibcoBRL, Gaithersburg, MD). The *D. melanogaster* Schneider II cell line was a gift from Dr. Allan Spradling (Carnegie Institution, Baltimore, MD). They were grown in Schneider’s *Drosophila* medium supplemented with 10% fetal bovine serum (FBS) and 50 μg/ml penicillin-streptomycin-glutamine in a 25°C ambient air incubator. HeLa cell stocks were grown in Delbecco’s Modified Eagles Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 50 μg/ml gentamycin in a 37°C, 5% CO₂ environment.

For transfection of HeLa cells, approximately 1 x 10⁵ cells were grown on 22 x 22 mm coverslips in six well culture plates (e.g. Falcon 3046) in DMEM at 37°C in a 5% CO₂ atmosphere for 12-18 hours. Transfection was by DNA-calcium phosphate precipitation using the N₁N-bis (2 – hydroxyethyl) – 2 – aminoethanesulfonic acid (BES) method of Chen and Okayama (1988). Ten μg of DNA per well was incubated with 0.25 M CaCl₂ and 1 x BBS for 20-30 minutes during which time a DNA precipitate forms. After adding the DNA precipitate, cells were incubated at 35°C in 3% CO₂ for 12 to 18
hours. Cells were then washed twice with 1 x phosphate buffered saline (PBS) and either fixed immediately in 2% paraformaldehyde in 0.6 x PBS, or recultured in DMEM with 10% FBS at 37°C in a 5% CO₂ atmosphere for an additional 24 hours prior to PBS washing and formaldehyde fixation. Transfection methods for Schneider II cells were identical to those methods just described except that they were maintained in D. melanogaster Schneider’s medium in a 25°C ambient air incubator.

**In vitro mRNA Production and Oocyte Injection**

Capped mRNAs encoding EGFP-tagged versions of DmNoppl40-RGG, DmNoppl40ΔRGG, and DmNoppl40 were synthesized using a T7 Message Machine kit from Ambion, Inc. (Austin, TX). The vectors pBluescript KS+/DmNoppl40-RGG, pBluescript KS+/DmNoppl40, and pBluescript/EGFP-DmNoppl40ΔRGG were linearized with Eco RI, Xho I, and Acc I respectively for *in vitro* run-off transcription reactions. Reactions and RNA purifications were performed according to Ambion’s (Austin, TX) instructions. Following transcription, RNAs were stored in ddH₂O at -70°C. Transcripts were injected into stage IV or V *X. laevis* oocytes using a Narishige microinjection apparatus. Following injection, the oocytes were cultured overnight at 18°C in OR2 medium (see Gall, 1998). The next day, oocyte nuclear contents were prepared for light microscopy as described by Gall (1998).

**Results**

**Exogenous Expression in D. melanogaster Schneider II Cells**

Full-length cDNAs encoding either DmNoppl40-RGG or DmNoppl40 were ligated into the mammalian expression vectors pEGFP-C3 and pDsRed1-C1 for the purpose of exogenous expression of GFP-tagged or RFP-tagged versions of the two proteins. A full length cDNA encoding *X. laevis* coilin was ligated into the same vectors.
Figure 4.1  *D. melanogaster* Schneider II cells transfected to express DmNopp140-RGG or DmNopp140. **Panels A and B** EGFP-DmNopp140-RGG localizes to the single prominent nucleolus of Schneider II cells (arrow in panels A and B). **Panels C and D** EGFP-DmNopp140 also localizes to nucleoli of Schneider II cells (arrow in panel C and D). No other organelles were labeled. Calibration bar is 40 μm for all panels.
in order to perform co-transfections with either of the *D. melanogaster* Nopp140 splice variants. Because the proteins in question are *D. melanogaster* specific, a common *D. melanogaster* cell line, Schneider II, was used for initial transient transfections.

Schneider II cells contain one prominent nucleolus. Both GFP-DmNopp140-RGG (Fig. 4.1A and B) and GFP-DmNopp140 (Fig. 4.1C and D) localized to the single nucleolus of *D. melanogaster* Schneider II cells. Transfection efficiencies were consistently low, but in the few transfected Schneider II cells, both proteins always localized to the nucleolus. Localization patterns between DmNopp140-RGG and DmNopp140 were indistinguishable, and no other nuclear bodies were labeled with either protein.

Rat Nopp140 not only localizes to nucleoli, but it also localizes to CBs. In fact, there is a physical interaction between rat Nopp140 and p80 coilin, the resident protein of CBs (Isaac et al., 1998). To determine if DmNopp140-RGG and/or DmNopp140 could localize to CBs in Schneider II cells, co-transfections with either GFP-DmNopp140-RGG or GFP-DmNopp140 and RFP-*X. laevis* coilin were performed (Fig. 4.2). *X. laevis* coilin is considered to be a homologue of p80 coilin (see Bellini, 2000), thus *X. laevis* coilin will be referred to as coilin throughout Chapter 4.

Exogenously expressed RFP-coilin localized to extra-nucleolar bodies that may be CBs (Fig. 4.2C and F). GFP-DmNopp140-RGG localized to nucleoli, but failed to associate with RFP-coilin within the putative CBs (Fig. 4.2B and C).

In cells expressing exogenous DmNopp140-RGG, however, there was a small but detectable fraction of RFP-coilin within the nucleoli (Fig. 4.2C). The RGG domain present within the carboxy-terminus of DmNopp140-RGG may prevent the protein from localizing to CBs, while still allowing an interaction with coilin. This specific interaction could cause the redistribution of coilin to the nucleoli of cells expressing DmNopp140-RGG.
Figure 4.2 *D. melanogaster* Schneider II cells transfected to co-express *X. laevis* coilin and either DmNopp140-RGG or DmNopp140. When GFP-DmNopp140-RGG was co-expressed with RFP-coilin Panels A-C DmNopp140-RGG localized to nucleoli (see Fig. 4.1), but it failed to localize with coilin in small extra-nucleolar bodies (panel C). Some RFP-coilin localized to the nucleoli (arrows in panel C). When GFP-DmNopp140 was co-expressed with RFP-coilin Panels D-F GFP-DmNopp140 localized to the single nucleolus (arrows in panels E and F) and to a single extra-nucleolar body (arrowhead in panels E and F). The small extra nucleolar body that labeled with GFP-DmNopp140 in panel E was brightly labeled with RFP-coilin in panel F. Calibration bar is 40 μm for all panels.
RGG and coilin. Conversely, GFP-DmNoppl40 localized to prominent nucleoli, and it co-localized with RFP-coilin within CBs (Fig. 4.2E and F). The co-localization of DmNopp140 and coilin in CBs agrees well with the findings of Isaac et al. (1998) who originally showed that rat Nopp140 localized to CBs.

A redistribution of coilin to nucleoli was not observed in Schneider II cells when co-expressed with DmNopp140 (FIG. 4.2G and F). Based upon the differences in localization patterns of coilin between Schneider II cells co-expressing DmNopp140-RGG or DmNopp140, the RGG domain may serve as a strong nucleolar localization signal or anchor that prevents DmNopp140-RGG from co-localizing with coilin within CBs in Schneider II cells.

**Exogenous Expression in X. laevis Oocytes**

The *X. laevis* oocyte nucleus, also called the germinal vesicle (GV), has a diameter of 300-400 μm. Its large size allows for hand isolation with jeweler’s forceps under a dissecting microscope. The GV contains large extra-chromosomal nucleoli, and thus the GV serves as an excellent system to study the sub-nucleolar localization of nucleolar proteins. To take advantage of this system, capped mRNAs encoding GFP-DmNopp140-RGG and GFP-DmNopp140 were synthesized and injected into stage IV - V *X. laevis* oocytes. Following injection, oocytes were incubated at 18°C overnight to allow for translation of the mRNAs and translocation of the protein from the cytoplasm to the nucleolus. The nuclear contents were then isolated and prepared for light microscopy according to Gall (1998). Before visualization, slides were stained with DAPI to mark the rDNA within the fibrillar centers (FC) of the multiple nucleoli. Both GFP-DmNopp140-RGG (Fig. 4.3A-C) and GFP-DmNopp140 (Fig 4.4E-G) localized to the region of the multiple nucleoli corresponding to the dense fibrillar component (DFC).
Figure 4.3 DmNopp140-RGG and DmNopp140 localize to the DFCs of *X. laevis* oocyte nucleoli. *X. laevis* oocytes were injected with synthetic transcripts that encoded either GFP-DmNopp140 (panels A-C) or GFP-DmNopp140 (panels D-F). Panels A and D are phase contrast images of a few of the approximately 1000 nucleoli from these oocyte nuclei. GFP-DmNopp140-RGG (panel B) and GFP-DmNopp140 (panel E) both localize within the DFC regions of the nucleoli, specifically within a sub-region immediately surrounding the FCs. The FCs contained detectable amounts of nucleolar DNA as demonstrated by DAPI staining (panels C and F). Some nucleoli typically contained multiple FCs and DFCs (panels C and F). Calibration bar is 20 μm for all panels.
immediately surrounding the FCs. Based upon the localization of DmNopp140 and coilin within putative CBs in Schneider II cells (Fig. 4.2D-F), DmNopp140 should also localize to the several C-snurposomes (CBs, see below) within X. laevis GV nuclei. In three separate injection trials, however, DmNopp140 never localized to X. laevis oocyte CBs that were easily identified by light microscopy (not shown).

**Expression of DmNopp140-RGG and DmNopp140 in HeLa Cells**

Mammalian expression vectors used to exogenously express GFP-tagged versions of DmNopp140-RGG and DmNopp140 and RFP-coilin in the Schneider II cells were also used to determine localization patterns of these proteins in HeLa cells. GFP-DmNopp140-RGG localized to HeLa cell nucleoli in a punctate staining pattern (Fig. 4.4A and B). In all HeLa cells transfected with only GFP-DmNopp140-RGG, the protein localized exclusively to nucleoli.

Exogenous expression of GFP-DmNopp140 caused HeLa cell nucleoli to segregate into phase-light and phase-dark regions. In some HeLa cells, the phase-light region surrounded the phase-dark region and extended well into the nucleoplasm (Fig. 4.4C and D). In other cases, the phase-light region appeared engulfed by the phase-dark region (Fig. 4.6A and B). There was no preference for a particular segregation pattern as both were well represented on any one cover slip examined. Whether expressed in relatively low or high amounts, GFP-DmNopp140 was always observed in the phase-light region of segregated nucleoli.

**Localization of Endogenous Fibrillarin and Nucleolin in HeLa Cells Expressing DmNopp140-RGG or DmNopp140**

To determine if exogenous expression of DmNopp140-RGG had any effect on the normal distribution of nucleolar proteins, HeLa cells expressing DmNopp140-RGG were
Figure 4.4 Expression of GFP-DmNopp140-RGG and GFP-DmNopp140 in HeLa cells. GFP-DmNopp140-RGG localizes to intact, phase-dark nucleoli in a punctate staining pattern Panels A and B Conversely, expression of GFP-DmNopp140 segregated nucleoli into phase-light and phase-dark regions Panels C and D GFP-DmNopp140 localized exclusively to the phase-light regions of segregated nucleoli. The phase-dark regions (arrows in panels C and D) were completely devoid of GFP-DmNopp140. Calibration bars are 20 μm for Panels A-D.
counter-stained with either anti-fibrillarin (Fig. 4.5C) or anti-nucleolin (Fig. 4.5F). Endogenous fibrillarin co-localized with GFP-DmNopp140-RGG in cells expressing moderate amounts of DmNopp140-RGG. Note that the punctate staining of DmNopp140-RGG in nucleoli is extinguished in HeLa cells expressing greater amounts of DmNopp140-RGG (arrow heads in Fig. 4.5B and C). Endogenous nucleolin also co-localized with GFP-DmNopp140-RGG within intact, phase-dark nucleoli (Fig. 4.5E and F). These results agree well with previous studies that show that Nopp140, fibrillarin, and nucleolin all reside within the DFC of nucleoli (Meier and Blobel, 1990; Snaar et al., 2000; Ginisty et al., 1999).

HeLa cells expressing DmNopp140 and displaying segregated nucleoli were subsequently stained with either anti-fibrillarin (Fig. 4.6C) or anti-nucleolin (Fig. 4.6F). Endogenous fibrillarin co-localized with GFP-DmNopp140 in the phase-light regions of segregated nucleoli (Fig 4.7A-C). This observation supports the finding that Nopp140 and fibrillarin both associate with snoRNPs and perhaps during pre-rRNA processing. Interestingly, while endogenous fibrillarin co-localized with DmNopp140 in the phase-light domains, all endogenous nucleolin localized within the phase-dark regions of segregated nucleoli (Fig. 4.6D-F). This suggests that nucleolin does not directly interact in vivo with either DmNopp140 or fibrillarin, although all three proteins normally associate with DFCs.

Localization of DmNopp140-RGG and DmNopp140 with X. laevis Coilin
Localization patterns for DmNopp140-RGG were dramatically different in HeLa cells that co-expressed coilin as compared to cells that expressed DmNopp140-RGG alone (compare Fig. 4.4A and B with Fig. 4.7A-C). In the presence of exogenous coilin, DmNopp140-RGG continued to localize to nucleoli in a punctate manner, but for the first
Figure 4.5 Endogenous fibrillarin and nucleolin co-localize with DmNopp140-RGG within nucleoli of HeLa cells. Panels A-C GFP-DmNopp140-RGG (panel B) co-localized with endogenous fibrillarin (panel C) that was detected with an anti-fibrillarin antibody and a rhodamine-conjugated secondary antibody. Arrowheads in panels A-C point to nucleoli that contain relatively large amounts of GFP-DmNopp140-RGG, while arrows in the same panels point to nucleoli that contain moderate amounts of GFP-DmNopp140-RGG and that maintain a punctate staining pattern. Panels D-F GFP-DmNopp140-RGG (panel E) co-localized with endogenous nucleolin (panel F) that was detected with an anti-nucleolin antibody and a rhodamine-conjugated secondary antibody. Arrows in panels D-F point to nuclear bodies that fail to stain with GFP-DmNopp140-RGG or anti-nucleolin. Conversely, all labeled structures evident in panels B and E correspond well to phase-dark nucleoli in phase-contrast micrographs A and D, respectively. Calibration bar is 20 μm for all panels.
Figure 4.6 Endogenous fibrillarin, but not endogenous nucleolin, co-localizes with GFP-DmNopp140 in phase-light regions of segregated HeLa cell nucleoli. Panels A-C GFP-DmNopp140 (panel B) co-localized with endogenous fibrillarin (panel C) in the phase-light regions of segregated nucleoli. An anti-fibrillarin antibody and a rhodamine-conjugated secondary antibody detected the endogenous fibrillarin (panel C). Panels D-F GFP-DmNopp140 (panel E) again localized to the phase-light regions, while endogenous nucleolin (panel F) localized to the phase-dark regions of segregated nucleoli. An anti-nucleolin antibody and a rhodamine-conjugated secondary antibody detected the endogenous nucleolin (panel F). Calibration bar is 20 μm for all panels.
DmNoppl40 co-localized with the coilin in several extra nucleolar bodies that may be CBs (Fig. 4.7C). RFP-coilin failed to localize to the nucleoli that appeared as black holes in Fig. 4.7C. This observation contrasts with the results obtained when GFP-tagged DmNopp140-RGG and RFP-coilin were co-expressed in D. melanogaster Schneider II cells. In the Schneider II cells, some coilin localized to nucleoli (Fig. 4.2B and C). This was not the case in HeLa cells co-expressing DmNopp140-RGG and coilin. The CBs that formed in these HeLa cells were larger than previously described for the exogenous expression of coilin (Sleeman et al., 1998). The size of the CBs in Fig. 4.7C was comparable to the size of CBs in HeLa cells expressing only GFP-X. laevis coilin (not shown). Therefore, the enlarged size of these CBs is likely due to expression of exogenous coilin.

When GFP-DmNopp140 and RFP-coilin were co-expressed in HeLa cells (Fig. 4.7D-F), the nucleoli were again segregated into phase-light and phase-dark regions. DmNopp140 again localized only to the phase-light regions of the segregated nucleoli (arrows in Fig. 4.7D-F). DmNopp140 and coilin co-localized in what may be very large CBs (arrowheads in Fig. 4.7D-F). Although these CBs (2 μm) are larger than the prototypical CBs (0.5 – 1 μm), enlarged CBs are also observed when coilin is expressed with DmNopp140-RGG (Fig. 4.7A-C). Coilin did not localize to either the phase-light or phase-dark regions of the segregated nucleoli (arrows in Fig. 4.7D-F) in cells co-expressing DmNopp140 and coilin.

Co-expression of DmNopp140-RGG and DmNopp140 in HeLa Cells

To determine if the two D. melanogaster splice variants would co-localize within intact or segregated nucleoli, GFP-DmNopp140 and RFP-DmNopp140-RGG were

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 4.7 HeLa cells transfected to co-express GFP-DmNopp140-RGG, GFP-DmNopp140, or *X. laevis* coilin. Panels A-C GFP-DmNopp140-RGG (panel B) was co-expressed with RFP-coilin (panel C). DmNopp140 localized to nucleoli and to several CBs, while coilin localized only to the CBs, but not to the nucleoli. Panels D-F GFP-DmNopp140 and RFP-coilin were co-expressed in HeLa cells. DmNopp140 again segregated nucleoli into phase-light (not marked) and phase-dark regions (arrows in panels D-F). DmNopp140 co-localized with exogenous coilin in CBs (arrowheads in panels D-F), but coilin failed to localize to the phase-light regions of the nucleoli. Calibration bar is 20 μm in all panels.
Figure 4.8 Co-expression of GFP-DmNopp140 and RFP-DmNopp140-RGG in transfected HeLa cells. Panels B, E, and H show expression of GFP-DmNopp140. Panels C, F, and I show expression of RFP-DmNopp140-RGG. Panels A-C GFP-DmNopp140 and RFP-DmNopp140-RGG were expressed in approximately equal amounts based upon fluorescence intensities. Both proteins localized to the nucleoli. One nucleolus in a transfected cell appeared partially segregated by phase contrast microscopy (lower arrow in panel A), but no more so than a nucleolus in a non-transfected cell (upper arrow in panel A). Panels D-F GFP-DmNopp140 (panel E) was over-expressed with respect to RFP-DmNopp140-RGG (panel F) based upon fluorescence intensities. Nucleoli appeared segregated by phase contrast microscopy (panel D), yet both proteins co-localized to the phase-light regions. Panels G-I RFP-DmNopp140-RGG (panel I) was overexpressed relative to GFP-DmNopp140 (panel H). Nucleoli appeared morphologically normal by phase contrast microscopy (panel G), yet both proteins co-localized. Calibration bar is 20 μm for all panels.
co-expressed in HeLa cells. Recall that DmNopp140-RGG, when expressed in HeLa cells, localized to intact phase-dark nucleoli. DmNopp140, when expressed in HeLa cells, caused the nucleoli to segregate into phase-light and phase-dark regions. DmNopp140 associated only with the phase-light regions (Fig. 4.4C and D).

When both DmNopp140-RGG and DmNopp140 were co-expressed in approximately equal amounts based upon fluorescence intensities, nucleoli remained relatively intact (Fig 4.8A-C); these nucleoli were not segregated into phase-light and phase-dark regions, and both *D. melanogaster* proteins localized to the intact nucleoli of HeLa cells. When more DmNopp140 was present relative to DmNopp140-RGG, nucleoli again appeared segregated (Fig 4.8D-F). In this case, phase-light regions were surrounded by phase-dark regions as previously observed. Surprisingly, both DmNopp140-RGG and DmNopp140 localized to the phase-light regions of these segregated nucleoli. When more DmNopp140-RGG was present relative to DmNopp140, nucleoli appeared morphologically normal and intact. Both *D. melanogaster* proteins localized to the intact, phase-dark nucleoli.

**Exogenous Expression of an RGG Truncation of DmNopp140-RGG in HeLa Cells**

The RGG domain of *X. laevis* nucleolin is necessary but not sufficient for its nucleolar localization (Heine et al, 1993). To determine if this is the case for DmNopp140-RGG, an RGG domain truncation expression vector was constructed. PCR primers were designed to amplify only amino acids 1-561 of DmNopp140-RGG. This eliminated the RGG domain as well as two additional tandem RGG motifs (amino acid residues 562-567) located upstream of the larger carboxy RGG domain. The deleted residues (562-688) are shown in italics in Fig 2.2A. This truncation is referred to as DmNopp140ΔRGG. When GFP-DmNopp140ΔRGG was expressed alone, the majority
Figure 4.9 Expression of the carboxy terminal truncation, DmNopp140ΔRGG, in transfected HeLa cells. **Panels A-C** GFP-DmNopp140ΔRGG was expressed on its own (panel B). Nucleoli were not labeled in most transfected cells (e.g. nucleus on the right in panel B). The truncated protein localized to small foci that may be endogenous CBs (arrows in panels B and C). The cells were counter-stained with anti-fibrillarin and a rhodamine-conjugated secondary antibody (panel C). Most of the endogenous fibrillarin remained associated with the nucleoli, but some fibrillarin co-localized with DmNopp140ΔRGG in the small nuclear bodies (arrows in panels B and C). **Panels D-F** GFP-DmNopp140ΔRGG localized to one large (arrow in panel E) and several small nuclear bodies that may be endogenous CBs. Nucleoli were again devoid of GFP-DmNopp140ΔRGG. These cells were counter-stained with anti-nucleolin and a rhodamine-conjugated secondary antibody (panel F). Nucleolin localized only to the phase-dark nucleoli. **Panels G-I** Transfected HeLa cells co-expressed GFP-DmNopp140ΔRGG (panel H) and RFP-coilin (panel I). Both proteins co-localized to enlarged CBs (not marked). Only GFP-DmNopp140ΔRGG localized to a phase light region (arrows in panels G-I) of a segregated nucleolus. Calibration bar is 20 μm.
of cells displayed intact nucleoli that were completely unlabeled. Instead, DmNopp140-ΔRGG localized to the nucleoplasm and to small nuclear bodies (Fig. 4.9B). In some cells the entire nucleoplasm was labeled except for the nucleoli. That appeared as dark holes within the nucleoplasm (not shown). In most cells expressing DmNopp140ΔRGG, the nucleoli appeared morphologically normal by phase contrast microscopy (Fig. 4.9A).

To determine if endogenous nucleolar proteins are redistributed by the presence of DmNopp140ΔRGG, HeLa cells expressing DmNopp140ΔRGG were stained with anti-fibrillarin (Fig. 4.9C). Most of the endogenous fibrillarin remained in the intact phase-dark nucleoli while some co-localized with DmNopp140ΔRGG in small extra nuclear bodies. These small nuclear bodies could be CBs (arrow in Fig. 4.9B and C). In some cells, the fibrillarin staining was somewhat more diffuse. This could be due to interactions between DmNopp140ΔRGG in the nucleoplasm and endogenous fibrillarin. In other cells (less than 5% of transfected cells), GFP-DmNopp140ΔRGG accumulated in large phase-light regions (Fig. 4.9D-F), similar to what was observed for DmNopp140 (Fig. 4.4C and D). This may be due to overexpression of DmNopp140ΔRGG in the cell in Fig. 4.9D-F. The same cell was counter-stained with anti-nucleolin (Fig. 4.9F). Again, endogenous nucleolin only associated with the phase-dark nucleoli, not the phase-light regions. In summary, DmNopp140ΔRGG labels the nucleoplasm and localizes to smaller extra nuclear bodies that may be CBs (see below). When DmNopp140ΔRGG is expressed at higher levels, however, it mimics full-length DmNopp140 by segregating nucleoli into phase-light and phase-dark regions.

In order to determine if DmNopp140ΔRGG could localize with coilin, HeLa cells were co-transfected with vectors that expressed GFP-DmNopp140ΔRGG and RFP-coilin. In cells expressing both DmNopp140ΔRGG and coilin, there was a dramatic shift in the
localization of DmNopp140ΔRGG. In these cells, DmNopp140ΔRGG and coilin co-localized within a few, relatively large CBs (Fig. 4.9G-I). There was only a slight nucleoplasmic signal, and small extra-nuclear bodies were evident. The CBs that did appear could not be detected by phase contrast microscopy (Fig. 4.9G). Most nucleoli were morphologically normal (intact), but some did segregate slightly into phase-light and phase-dark regions (arrows Fig. 4.9H-I). Some DmNopp140ΔRGG localized to the phase-light regions of these segregated nucleoli, while coilin localized only to the periphery of the nucleolar regions, but not within them. The localization pattern of DmNopp140ΔRGG therefore changes dramatically when co-expressed with coilin. Its localization pattern is much like that of DmNopp140-RGG when expressed with coilin in that it accumulates in relatively large CBs when coilin is co-expressed. This observation is consistent with earlier observations by Isaac et al. (1998) that showed a direct interaction between Nopp140 and coilin. Here, however, the carboxy tail of DmNopp140-RGG is deleted. Also, DmNopp140ΔRGG co-localizes with fibrillarin, but not coilin in phase-light regions of segregated nucleoli (Fig. 4.9I). DmNopp140ΔRGG did not co-localize with endogenous nucleolin in phase-dark regions (Fig. 4.9G-I).

**Discussion**

**Exogenous Expression of DmNopp140-RGG, DmNopp140, and *X. laevis* Coilin in *D. melanogaster* Schneider II Cells**

GFP-DmNopp140 and GFP-DmNopp140-RGG localized to the single prominent nucleolus of *D. melanogaster* Schneider II cells. These observations were expected due to comparisons made between the amino acid sequences of both *D. melanogaster* Nopp140 variants and vertebrate Nopp140. DmNopp140-RGG and DmNopp140 share an alternating acidic/basic domain with vertebrate Nopp140, nucleolin, B23, and yeast
NSR1 and NPI46. Also, DmNopp140-RGG shares an RGG domain with the nucleolar proteins nucleolin, fibrillarin, and B23 to name a few.

GFP-DmNopp140 localized to nucleoli (Fig. 4.1C and D), and it co-localized with RFP-coilin in CBs (Fig. 4.2D-F). The co-localization of GFP-DmNopp140 and RFP-coilin was expected (Isaac et al., 1998), and the result presented here strengthens the argument that DmNopp140 is the true homologue of rat Nopp140. GFP-DmNopp140-RGG localized to nucleoli of Schneider II cells as expected (Fig. 4.1A and B). What was not expected, however, was the localization patterns of GFP-DmNopp140-RGG when co-expressed with RFP-coilin (Fig. 4.2A-C). In this case, exogenously expressed coilin localized to the nucleolus and the CBs. The effects of the RGG domain on DmNopp140-RGG's localization may be two fold. First, the RGG tail may prevent GFP-DmNopp140-RGG from localizing to CBs in Schneider II cells. Second, the presence of the RGG tail may cause some of the RFP-coilin that is co-expressed to localize to the nucleolus by maintaining an interaction with DmNopp140-RGG. From these observations, I propose that the RGG tail of DmNopp140-RGG serves as a strong nucleolar retention signal or anchor that prevents DmNopp140-RGG from accumulating within or shuttling through CBs in Schneider II cells.

The appearance of coilin within nucleoli of Schneider II cells exogenously expressing both DmNopp140-RGG and coilin was surprising. Coilin has been observed in and around nucleoli, but only in extraordinary circumstances. For example, when cells are treated with actinomycin D, a general inhibitor of transcription, coilin is found clustered around nucleoli (Carmo-Fonseca et al., 1992). Treatment of somatic cells with the specific serine/threonine phosphatase inhibitor, okadaic acid, also redistributes coilin and splicing snRNPs to discrete subnucleolar domains (Lyon et al., 1997; Sleeman et al,
Finally, coilin is observed in the nucleolus when serine 202 is substituted with an aspartic acid, which mimics constitutive phosphorylation (Lyon et al., 1997). Interestingly, in most of these cases, coilin associates with a defined structure within nucleoli, while in Schneider II cells expressing DmNopp140-RGG and coilin, the coilin is diffusely spread throughout the nucleus.

Unlike the results obtained with co-expression of DmNopp140-RGG and coilin, co-localizations of DmNopp140 and coilin in CB-like structures suggest that Schneider II cells and thus other D. melanogaster cells could contain CBs. This is important because previous attempts failed to stain CBs within D. melanogaster cells with antisera against p80 coilin and X. laevis coilin (Dr. Joe Gall, personal communication).

**Exogenous Expression of DmNopp140-RGG and DmNopp140 in X. laevis Oocytes**

DmNopp140-RGG and DmNopp140 localized to the DFC of the multiple nucleoli of X. laevis oocytes. The DFC is a region that is enriched in endogenous fibrillarin. The localization of both D. melanogaster proteins to the DFC is consistent with the observation that vertebrate Nopp140 associates with fibrillarin in box C/D snoRNPs (Yang et al., 2000). Although a direct physical interaction between these two proteins has not been determined, antibodies directed against Nopp140 can immunoprecipitate snoRNP complexes containing fibrillarin. The localizations of the D. melanogaster Nopp140 splice variants to the DFC immediately surrounding the FC (DFC borders) suggests that these proteins are somehow involved in pre-rRNA transcription and/or processing (see Chapter 1). Shah et al. (1996) showed that fibrillarin enriches within this same border region of the oocyte nuclei.

Surprisingly, the multiple nucleoli of the oocytes, but not the CBs, were labeled. The oocyte CBs (formerly known as spheres or C snurposomes, see Gall, 2000) should
have been labeled, since they are known to contain Nopp140 (Gall et al, 1999). Three separate injection trials were performed, and in all three, no other nuclear bodies besides nucleoli were labeled with GFP-DmNopp140-RGG or GFP-DmNopp140. The possibility remains that the D. melanogaster Nopp140 variants are significantly different in sequence from the X. laevis Nopp140 homologue (Cairns and McStay, 1995) that they are excluded from the oocyte CBs. The alternating acidic and basic domain of Nopp140 is required for CB localization (Isaac et al, 1998). The X. laevis Nopp140 homologue contains 18 blocks of acidic residues alternating with 17 blocks of basic residues (Cairns and McStay, 1995). Both D. melanogaster Nopp140 variants contain 15 blocks of acidic residues interspersed with 14 blocks of basic residues (see Fig. 1.1A and B). These differences could explain the failure of the D. melanogaster Nopp140 variants to localize to oocyte CBs.

**Exogenous Expression of DmNopp140-RGG, DmNopp140, and X. laevis Coillin in HeLa Cells**

DmNopp140-RGG localized to CBs in HeLa cells only when co-expressed with coillin. This brings up the question: Does DmNopp140-RGG localize to endogenous CBs? I have observed countless transfected cells and conclude that DmNopp140-RGG only localizes to enlarged CBs when co-expressed with exogenous coillin. When DmNopp140-RGG is solely expressed in HeLa cells, the protein does not localize to smaller endogenous CBs. I propose that DmNopp140-RGG does not normally associate with endogenous coillin in CBs, but when both proteins are over-expressed in a cell, interaction between the two proteins causes DmNopp140-RGG to localize to enlarged CBs. Rat Nopp140 interacts with p80 coillin (Isaac et al., 1998). This interaction could occur in regions that are common to both DmNopp140-RGG and DmNopp140.
These enlarged CBs contain tri-methyl capped snRNAs (not shown). Therefore they are not merely aggregations of D. melanogaster Nopp140 and coilin. These enlarged CBs could be due to the fusing of two or more CBs, or enlarged CBs could be due to aggregations of DmNopp140-RGG and coilin without other known CB components. In order to determine if enlarged CBs observed when DmNopp140-RGG and coilin are co-expressed are true CBs, experiments must be performed to localize other known CB components to these structures.

These results differ from those in Schneider II cells that co-expressed GFP-DmNopp140-RGG and RFP-coilin. In Schneider II cells that exogenously expressed both DmNopp140-RGG and coilin, some coilin was redistributed to nucleoli in a diffuse manner (Fig. 4.1A-C). Coilin localized to only CBs and not to nucleoli in Schneider II cells expressing only coilin (data not shown).

DmNopp140-RGG co-localized with endogenous fibrillarin (Fig. 4.7A-C) and nucleolin (Fig. 4.7D-F) in HeLa cells. Here, DmNopp140-RGG localized to regions corresponding to the DFC, a region enriched in fibrillarin (Fig. 4.3A-C). These results correlate with results obtained when synthetic mRNAs encoding GFP-DmNopp140-RGG were injected into X. laevis oocytes. Mammalian Nopp140 interacts with fibrillarin in box C/D snoRNP complexes (Yang et al, 2000). Although no physical interactions between Nopp140 and nucleolin have been established, endogenous nucleolin co-localized well with exogenously expressed DmNopp140-RGG (Fig. 4.7D-F) in intact phase-dark nucleoli. This was expected because both proteins are known to reside within the DFCs of nucleoli.

Smaller, extra-nuclear bodies (perhaps CBs) were also observed in these cells by phase contrast microscopy. Neither exogenously expressed DmNopp140-RGG nor
endogenous nucleolin localized to these bodies. This is significant because in those cells over expressing GFP-DmNopp140-RGG, the nucleoplasm is lightly but uniformly labeled, but DmNopp140-RGG only accumulates in the phase dark nucleoli. Because CBs within somatic cells do not contain nucleolin, these results suggest that DmNopp140-RGG localizes only to nucleoli and not CBs when exogenously expressed on its own in HeLa cells without the expression of exogenous coilin.

In summary, DmNopp140-RGG localizes to phase-dark nucleoli of HeLa cells in a punctate manner. This punctate staining suggests subdomain structure within the DFC of HeLa cell nucleoli. These subcompartments may be analogous to the DFC/FC border regions observed in the large nucleoli of *X.* *laevis* oocytes. Co-localizations of DmNopp140-RGG with endogenous fibrillarin and nucleolin within intact nucleoli suggest that over expression of DmNopp140-RGG has no apparent effect on the normal nucleolar morphology of HeLa cells. Only when co-expressed with exogenous coilin does DmNopp140-RGG localize to extra-nucleolar bodies.

DmNopp140 and coilin co-localize to extra nucleolar bodies that may be very large CBs. Enlarged CBs are also observed when an amino-terminal truncation of the survival of motor neurons (SMN) protein (SMNΔN27) is over-expressed in HeLa cells (Pellizzoni et al., 1998). The SMN proteins along with SIP1 are CB proteins involved in the biogenesis of snRNPs (Liu and Dreyfuss, 1996; Liu et al., 1997). Over expression of SMNΔN27 displays dominant negative effects by redistributing snRNAs and Sm proteins to large nuclear (and cytoplasmic) CBs (Pellizzoni et al., 1998).

**Segregation of HeLa Cell Nucleoli**

Expression of DmNopp140 in HeLa cells causes nucleoli to segregate into phase-light and phase-dark regions (Fig. 4.4C and D). DmNopp140 and endogenous fibrillarin
localize to the phase-light regions, while endogenous nucleolin remains in the phase-dark regions (see Fig. 4.6). Structures that may be related to these segregated nucleoli of Figure 4.5 have been detected previously. In one study, specific Nopp140 deletion constructs containing the central repeat domain caused formation of phase-dense structures called Nopp-R induced rings (R-rings) (Isaac et al, 1998). These R-rings ranged in size from 0.5-5.0 μm and varied between 2-5 per cell. The R-rings contained fibrillarin, NAP57, UBF, and in part RNA Pol I itself. Interestingly, the phase-dark nucleolus remained intact on a light microscopic level despite the absence of these major nucleolar proteins. Nucleolin and B23, however, remained in the phase-dark nucleoli. In addition, a subset of p80 coilin localized to R-rings, but the Sm antigens, known to also localize to CBs, were not found in the R-rings. One Nopp140 deletion, HA-NoppΔC, localized particularly within the cytoplasm, and it caused formation of R-rings within the cytoplasm. Therefore, both R-rings and segregated nucleoli can recruit subsets of nucleolar proteins while endogenous nucleolin is left behind in phase-dark nucleoli.

Although R-rings and segregated nucleoli are similar in some ways, differences are readily apparent. First, R-rings are formed by the expression of Nopp140 deletions, specifically those containing the central repeat domain. The segregated nucleoli shown in Figure 4.5 are formed by the expression of a full-length DmNopp140 construct. Second, the phase-light region remains attached to the phase-dark region of the segregated nucleoli. This is in contrast to the R-rings which can form external to nucleoli and even in the cytoplasm.

In another example, HeLa cells expressing a carboxyl truncation of human Nopp140 (hNopp140N382), displayed crescent shaped nucleoli. In this case Chen et al. (1999) showed that these crescent shaped nucleoli contain endogenous hNopp140, RNA
Pol I, and fibrillarin. Endogenous nucleolin on the other hand, redistributed from nucleoli to the nucleoplasm. The authors reasoned that these results were analogous to those obtained when cells were treated with actinomycin D (a transcription inhibitor). The authors went on to conclude that rRNA transcription was shut down by over expression of hNopp140N382 through its interaction with RPA194, the large subunit of RNA Pol I. Much of their work has suggested that Nopp140 is a transcription factor.

The segregated nucleoli in HeLa cells expressing DmNopp140 may allow us to explore possible molecular associations between nucleolar components in vivo. This phenomenon could become particularly valuable if the segregated nucleoli in question are still transcriptionally active. It is fairly well accepted that the bulk of the tandem rDNA genes reside within the FC, and that rDNA transcription occurs at the border between the FC and DFC. The segregation of nucleoli in HeLa cells expressing DmNopp140 may allow for better characterization of transcription sites within nucleoli, and where the nucleolar proteins involved in this process reside with respect to transcription.

Fibrillarin associates with the U3 snoRNA, and it is ubiquitous in all box C/D snoRNPs. Nucleolin also associates with the U3 snoRNP for cleavage within the 5' ETS of pre-rRNA (Ginisty et al, 1998). Both fibrillarin and nucleolin localize to the DFC of nucleoli. Taken together, these data would suggest that these proteins interact in vivo. The segregation of nucleoli where endogenous fibrillarin resides within the phase-light regions and where endogenous nucleolin resides in the phase-dark regions would suggest otherwise.

Rat Nopp140 is a component of both box C/D and H/ACA snoRNPs while fibrillarin is reserved for only box C/D snoRNPs. Since DmNopp140 and fibrillarin both localize only to the phase-light regions of segregated nucleoli, I predict that all
components of snoRNPs (i.e. NAP57, NAP65) will associate with phase-light regions of segregated nucleoli. Nucleolin remains associated with the phase-dark regions of segregated nucleoli. Therefore, I predict that nucleolin and B23 will associate with phase-dark material based upon the observation of Isaac et al (1998), and based upon the fact that nucleolin and B23 are known to associate (Liu and Yung, 1999). Nucleolin and B23 also remained associated with phase-dark nucleoli during the production of R-rings (Isaac et al., 1998). Thus, the segregation of nucleoli into phase-light and phase-dark regions may represent the segregation of nucleoli into two separate domains each with different sets of rRNA processing functions. It will be interesting to determine what other components reside within the phase-light and phase-dark regions of segregated nucleoli expressing DmNopp140.

Co-expression of DmNopp140-RGG and DmNopp140 in HeLa Cells

When DmNopp140-RGG and DmNopp140 were expressed in equal amounts in HeLa cells, both proteins localized to relatively intact nucleoli (Fig. 4.8A-C). When DmNopp140 was expressed in greater amounts relative to DmNopp140-RGG, the nucleoli segregated into phase-light and phase-dark regions (Fig. 4.8D-F). In this case, DmNopp140-RGG co-localized with DmNopp140 in the phase-light regions. In HeLa cells where more DmNopp140-RGG was expressed relative to DmNopp140, both proteins localized to intact, phase-dark nucleoli (Fig. 4.8G-I). Therefore, a titering effect seems to be at work, but more importantly, the results are consistent with an interaction between the two proteins.

It has not yet been determined whether both D. melanogaster Nopp140 variants are present within the same D. melanogaster cell at any one time. Results presented here suggest that when and if this event occurred, there may be an in vivo interaction between
the two proteins. In other words when DmNopp140 is in over-abundance, it can sequester DmNopp140-RGG. Conversely, when DmNopp140-RGG is in over-abundance, it can sequester DmNopp140.

**Exogenous Expression of DmNopp140ΔRGG**

DmNopp140-RGG, like nucleolin, contains a large RGG domain near its carboxy terminus. Several labs have determined that the RGG domain of nucleolin is necessary, but not sufficient for its nucleolar localization (Meßmer and Dreyer, 1993; Schmidt-Zachman and Nigg, 1993; Heine et al., 1993; Creancier et al., 1993).

DmNopp140ΔRGG, like a carboxy nucleolin truncation, failed to associate with nucleoli in HeLa cells (Fig. 4.9A-C). The same results were obtained in this dissertation research using CHO cells (data not shown). DmNopp140-RGG re-distributes to the nucleoplasm and to small extra nuclear bodies that are probably related to CBs (Fig 4.10B). Thus, the RGG domain of DmNopp140-RGG is also necessary for its nucleolar localization. In Schneider II cells, the percentage of cells in which DmNopp140ΔRGG localized to nucleoli was more variable. In some Schneider II cells, DmNopp140ΔRGG localized to nucleoli while in other cells, the truncated protein did not localize to nucleoli at all (not shown). Cross species differences may explain why DmNopp140ΔRGG can localize to nucleoli in Schneider II cells much better than to nucleoli in HeLa or CHO cells.

The nucleoplasmic localization of DmNopp140ΔRGG contrasts sharply with the nucleolar localization of a rat Nopp140 carboxy terminal truncation, NoppΔC (Isaac et al., 1998). The rat Nopp140ΔC was truncated such that the carboxy terminal end of NoppΔC ends with the central repeat domain. The DmNopp140-RGG truncation, on the other hand, left behind ten residues that are unique to DmNopp140-RGG such that these ten residues were the carboxy terminus of the protein. Unlike DmNopp140ΔRGG, NoppΔC
localized to intact nucleoli, CBs, and to R-rings. The differences in localization patterns between NoppΔC and DmNopp140ΔRGG may be due to the ten residues left behind on the carboxy terminus of DmNopp140ΔRGG.

Finally, the arginine residues within the RGG domains of nucleolin and fibrillarin are asymmetrically dimethylated (N°, N°-dimethylarginine). Methylation does not change the overall charge of the RGG domain, but the addition of these bulky side chains may modulate protein/protein interactions (Tao and Frankel, 1992; Liu and Dreyfuss, 1995; Friesen et al., 2001), or even its nucleo-cytoplasmic shuttling (Shen et al., 1998). The SMN protein binds preferentially to the dimethylarginine-modified RGG domains of SmD1 and SmD2, two of the seven Sm proteins (Friesen et al., 2001). The methylation of CHO nucleolin is under investigation in our lab. I predict that the RGG tail of DmNopp140-RGG will be methylated and that this unique post translational modification will differentially regulate DmNopp140-RGG function and associations, specifically protein-protein interactions, versus those of DmNopp140.
Chapter 5

Conclusions and Future Directions

The nucleolar phosphoprotein of 140kDa (Noppl40) was first described in the late 1980's as a protein that specifically recognizes the wild type (but not mutant) NLS within the SV-40 large T antigen (Yamasaki et al., 1989). Noppl40 also interacts with synthetic NLS peptides of nucleoplasmin and the adenovirus EIA protein (Yamasaki et al., 1989). Rat Noppl40 shuttles on curvilinear tracks between the nucleolus and cytoplasm (Meier and Blobel, 1992). These results suggest that Noppl40 may shuttle ribosomal subunits between the nucleolus and cytoplasm.

Noppl40 homologues have been described in X. laevis (Cairns and McStay, 1995), human (Yeh et al., 1995), yeast (Meier, 1996), and now D. melanogaster (this dissertation). Noppl40 functions as a transcription factor (Chen et al., 1997) and interacts directly with both casein kinase II (CKII) (Li et al., 1997) and the largest subunit of RNA Pol I, RPA194 (Chen et al., 1999). Two of the more interesting functions ascribed to Noppl40 are that it functions as a molecular link between the nucleolus and Cajal bodies (CBs) (Isaac et al., 1998) and it is a component of mammalian box H/ACA and box C/D snoRNPs (Yang et al., 2000). These reports suggest that Noppl40 functions in the transport of snoRNPs from nucleus to nucleolus where they function in the site-specific pseudouridylation and 2'-O-methylation of rRNA (reviewed in Tollervey and Kiss, 1997; Weinstein and Steitz, 1999).

This dissertation is the first report of a cDNA encoding Noppl40 from D. melanogaster. While non-ribosomal nucleolar proteins and the nucleolus in general have been studied extensively in other eukaryotes, information concerning these subjects in D. melanogaster has lagged behind. Very little is known about non-ribosomal nucleolar...
proteins and their interactions within nucleoli and CBs in \textit{D. melanogaster}. In fact, there have been no reports of any of the major non-ribosomal nucleolar proteins (nucleolin, B23, fibrillarin) in \textit{D. melanogaster}.

The gene encoding \textit{D. melanogaster} Nopp140 gives rise to two mRNAs as a result of alternative mRNA splicing. One of these mRNAs encodes a protein of 654 amino acid residues and is made up of four exons. The first two exons are made up of nucleotides 88-186 and 389-1942 and encode amino acid residues 1-33 and 34-551, respectively. The second two exons include nucleotides 2184-2342 and 3096-3248 which encode amino acid residues 552-604 and 605-654, respectively. This protein, DmNopp140, consists of two signature domains also present in mammalian Nopp140. These domains are a central domain consisting of alternating acidic and basic repeats, and a carboxy domain that is evolutionarily conserved in Nopp140 homologues from all species.

Three exons constitute the second splice variant of \textit{D. melanogaster} Nopp140. Nucleotides 88-186, 389-1942, and 2580-2993 encode amino acid residues 1-33, 34-551, and 552-688, respectively. Thus, the two spliced mRNAs are identical in their first two exons that encode amino acid residues 1-551. Due to alternative mRNA splicing, however, the two transcripts contain mutually exclusive exons that encode totally different carboxy termini. Amino acids 552-668 of DmNopp140-RGG encode a carboxy domain that contains an arginine and glycine rich domain made of arginine-glycine-glycine (RGG) repeats. This RGG domain is similar to the RGG domain present in vertebrate nucleolin and fibrillarin. It is this difference in carboxy termini that confers different localization patterns within \textit{D. melanogaster} Schneider II cells and HeLa cells (see Chapter 4).
DmNopp140 may perform the same functions in *D. melanogaster* as Nopp140 does in other species. These include, but are not restricted to, the putative shuttling of box H/ACA and box C/D snoRNPs, and the putative shuttling of ribosomal subunits. DmNopp140-RGG on the other hand may perform functions normally reserved for vertebrate nucleolin in *Drosophila* through its similar RGG domain. The RGG domain in DmNopp140-RGG may interact with nucleolar proteins or RNAs that typically associate with nucleolin. The arginine residues within the RGG domains of nucleolin and fibrillarin are asymmetrically dimethylated (N°, N°- dimethylarginine). The purpose of this post-translational modification is not known. Methylation does not change the overall charge of the RGG domain, but the addition of these bulky side chains may modulate protein/protein or protein/nucleic acid interactions (Tao and Frankel, 1992; Liu and Dreyfuss, 1995; Friesen et al., 2001), or even its nucleo-cytoplasmic shuttling (Shen et al., 1998).

CKII phosphorylates rat Nopp140 to an unusually high degree (Meier, 1996). This is not surprising based upon the 49 putative CKII phosphorylation sites present in rat Nopp140. A CKII consensus phosphorylation site consists of a serine or threonine acceptor site and an acidic amino acid three residues away on its carboxy-terminal side. Every acidic amino acid residue near the target serine improves the acceptor site (Marin et al, 1986; Kuenzel et al, 1987). Phosphorylation of the target serine within a Nopp140 acidic repeat causes the preceding serine to become a better substrate for CKII. This dissertation describes the labeling of *E. coli* expressed DmNopp140-RGG with gamma-labeled [³²P] – ATP in an in vitro phosphorylation reaction. DmNopp140-RGG contains 64 putative CKII phosphorylation sites, and like rat Nopp140, DmNopp140-RGG is labeled by CKII with gamma-labeled [³²P] – ATP to a high degree (Fig. 3.1). Based upon
the sequence similarities between DmNopp140-RGG and DmNopp140, DmNopp140 most probably is a CKII substrate as well.

Examination of the amino acid sequence of DmNopp140-RGG reveals eight potential p34<sup>cdk2/cyclin B</sup> (MPF) phosphorylation sites (Fig. 2.2A). *E. coli* expressed DmNopp140-RGG can also serve as a substrate for MPF in an *in vitro* phosphorylation reaction (Fig. 3.1). DmNopp140 may also serve as a substrate for MPF, based upon sequence similarities (Fig. 2.2B). MPF phosphorylation of Nopp140 may regulate nucleolar disassembly during prophase, while its dephosphorylation may regulate nucleogenesis during telophase (Pai et al., 1995).

This dissertation shows that DmNopp140-RGG is a substrate for CKII and MPF in *in vitro* phosphorylation reactions. In addition, close examination of the amino acid sequence of DmNopp140 reveals a putative cAMP dependent protein kinase phosphorylation site not present in the primary amino acid sequence of DmNopp140-RGG. This suggests that DmNopp140 may contain another level of regulation by signal transduction cascades that DmNopp140-RGG does not.

The two splice variants of *D. melanogaster* Nopp140 differ in their carboxy terminal ends. DmNopp140 is the bona fide *D. melanogaster* Nopp140 homologue in that it contains a carboxy terminus that is 58% identical to that of rat Nopp140 (Fig. 2.4A). DmNopp140-RGG contains an RGG domain at its carboxy terminus, and is thus, unique. GFP-tagged versions of both splice variants localize to the single prominent nucleoli of Schneider II cells (Fig 4.1). When RFP-coilin is co-expressed with GFP-DmNopp140 in Schneider II cells, the two proteins co-localize in small extra nucleolar nuclear bodies that may be CBs (Fig. 4.2D-G). Conversely, when RFP-coilin is co-expressed with GFP-DmNopp140-RGG in Schneider II cells, no co-localization is
observed. Interestingly, when DmNopp140-RGG is co-expressed with coilin in Schneider II cells, coilin localizes to the nucleoli (Fig. 4.2A-C). Perhaps the RGG domain of DmNopp140-RGG acts as a strong nucleolar localization or retention signal that reduces the amount of shuttling from the nucleolus to CBs in Schneider II cells. Coilin could redistribute to the nucleolus by maintaining an interaction with DmNopp140-RGG. It will be interesting to investigate why coilin localizes to the nucleolus in this case since coilin only redistributes to the nucleolus in special circumstances (see Chapter 1, Chapter 4).

Expression of synthetic mRNAs encoding GFP-tagged versions of DmNopp140 and DmNopp140-RGG show that these proteins localize to regions in the DFC of the multiple nucleoli immediately surrounding the FC (Fig. 4.3). The localizations of the D. melanogaster Nopp140 splice variants to the DFC immediately surrounding the FC (DFC borders) suggests that these proteins are directly involved in pre-rRNA transcription and/or processing (see Chapter 1). Shah et al. (1996) showed that fibrillarin enriches within this same border region of the oocyte nuclei.

It was suspected that DmNopp140 would localize to CBs (or C-snurposomes) of nuclear preps of X. laevis oocytes. This was not the case. Neither DmNopp140 nor DmNopp140-RGG localized to the CBs in this system. It is possible that slight differences between the sequences of the D. melanogaster Nopp140 variants and xNopp180 (X. laevis homologue of Nopp140) exclude the D. melanogaster proteins from the oocyte CBs (Cairns and McStay, 1995).

DmNopp140-RGG localized to intact nucleoli of HeLa cells in a punctate manner (Fig. 4.4A and B). When GFP-DmNopp140-RGG was co-expressed with RFP-coilin, DmNopp140-RGG co-localized with coilin in enlarged CBs (Fig. 4.7A and B).
determined by immunolabeling that both endogenous fibrillarin and nucleolin co-localize with DmNopp140-RGG in intact phase-dark nucleoli (Fig. 4.5).

Expression of DmNopp140 in HeLa cells causes nucleoli to segregate into phase-light and phase-dark regions (Fig. 4.4C and D). Endogenous fibrillarin localized with DmNopp140 in the phase-light regions of segregated nucleoli (Fig. 4.6A-C). Endogenous nucleolin localized to the phase-dark regions of segregated nucleoli of HeLa cells expressing DmNopp140 (Fig. 4.6D-F). The segregation of nucleoli into phase-light and phase-dark regions may reflect the segregation of nucleoli into separate functional domains. In the future it will be necessary to determine if these segregated nucleoli are still transcriptionally active. If these segregated nucleoli are indeed transcriptionally active, they could become valuable in determining where transcription takes place within nucleoli, and where the nucleolar proteins involved in this process occur with respect to transcription. Like DmNopp140-RGG, DmNopp140 localized to what may be enlarged CBs in HeLa cells when co-expressed with coilin (Fig. 4.7D-F).

Co-expression of DmNopp140 and DmNopp140-RGG in HeLa cells suggests that the extent to which nucleoli segregate is proportional to the amount of DmNopp140 that is present within the cell. In HeLa cells where more DmNopp140-RGG was expressed relative to DmNopp140, DmNopp140 localized to phase-dark nucleoli (Fig. 4.8G-I). Conversely, when DmNopp140 was in higher concentrations than DmNopp140-RGG, DmNopp140-RGG localized to phase-light regions of segregated nucleoli (Fig. 4.8D-F). These results are consistent with the possibility that the D. melanogaster Nopp140 splice variants may interact in vivo.

A carboxy truncation of DmNopp140-RGG, DmNopp140ΔRGG, did not localize to intact nucleoli, but labeled the entire nucleoplasm (Fig 4.9B). DmNopp140ΔRGG did,
however, localize to small extra-nuclear bodies that may be CBs (Fig 4.9B). Endogenous
fibrillarin and nucleolin localized to intact nucleoli, suggesting that nucleoli remained
morphologically normal (Fig. 4.9C and F). These results suggest that like nucleolin, the
RGG domain of DmNopp140-RGG is necessary for its nucleolar localization, but not for
its localization in CBs.

The most obvious question related to this project is which cell types of D. melanogaster express the splice variants. The differences in the carboxy termini of the
two D. melanogaster Nopp140 splice variants should be able to help answer this question
on two levels. First, RNA probes made from cDNAs encoding the different carboxy
termi should allow for in situ hybridizations to whole mount egg chambers or embryos.
This technique can be used to detect mRNAs encoding either DmNopp140 or
DmNopp140-RGG. Antibodies directed against short peptides designed from the
different carboxy termini of DmNopp140 and DmNopp140-RGG should allow detection
of endogenous protein. Immunostaining of whole mount or sectioned egg chambers
and/or embryos may show where these proteins are concentrated in these structures. In
addition, antibodies directed against the carboxy termini of the D. melanogaster Nopp140
variants could be used for immunoprecipitation reactions to determine the proteins that
interact with DmNopp140 and DmNopp140-RGG. DmNopp140 most likely would
interact with those proteins that the typical rat Nopp140 does (Nop60B, NAP65, coilin),
while DmNopp140-RGG would interact with proteins that normally interact with
nucleolin through its RGG domain.

Another possible project is to track the fates of DmNopp140 and DmNopp140-
RGG through oogenesis and embryogenesis by transformation of D. melanogaster. In
order to view DmNopp140 and DmNopp140-RGG in living tissues in real time, the two
proteins will be expressed as GFP fusions using P-element-mediated transformation. The P-element vector, \textit{pCaSpeR-hs-act} (GenBank Acc. No. U60735), will be used that contains the \textit{white} \textsuperscript{+} gene, the hsp70m heat shock promoter, and a poly[A] sequence from a cytoplasmic actin gene (act5c), all flanked by P-element ends (Spradling, 1986; Pirrotta, 1988; Thummel et al., 1988). This promoter is active in most stages of \textit{D. melanogaster} development and in the adult ovary (Fyrberg, 1983). The cDNAs encoding GFP-tagged versions of DmNopp140-RGG and DmNopp140 have been subcloned into the \textit{pCaSpeR-hs-act} vector. The helper plasmid, pUCHA2-3, encoding an active transposase, will be co-injected with the recombinant P-element vectors into the posterior of staged embryos of \textit{w} \textsuperscript{1118} females. Adult transformants will be identified by screening for restoration of wild-type eye pigmentation. Transposon insertion will be localized by in situ hybridizations of specific 3' cDNA subclones (unique to either DmNopp140 or DmNopp140-RGG) to the polytene chromosomes of \textit{D. melanogaster} salivary glands (Pardue, 1994). One advantage of older egg chambers (stage 10b and older) is that they continue to mature autonomously in culture. Once we see GFP-DmNopp140 or GFP-DmNopp140-RGG in nucleoli of cultured stage 10b egg chambers after P-element transformation, we can follow in real time the flow of nurse cell cytoplasm through the ring canals and into the oocyte as the nurse cells enter apoptosis. Expression of GFP-DmNopp140 or GFP-Nopp140-RGG by P-element transformation will also allow us to follow nucleogenesis in blastoderm cells, again in real time.

The activities of non-ribosomal nucleolar proteins are now understood to be important for the normal functions of both nucleoli and CBs. Although these proteins have been studied extensively in other eukaryotes, knowledge of non-ribosomal nucleolar proteins in \textit{D. melanogaster} lags far behind. The discovery of two splice variants of
Nopp140 in *D. melanogaster* may serve as a link to describing other non-ribosomal nucleolar proteins in *D. melanogaster*. The molecular cytology and genetic tools available in *D. melanogaster* should serve to greatly advance the knowledge of Nopp140 and non-ribosomal nucleolar proteins in general.
Literature Cited


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


interphase PNBs show similar characteristics as those typically observed at telophase of mitosis in untreated cells. *Chromosoma* **105**, 418-430.


Appendix A: Primer Locations

Compilation of sequence from lambda clones and LD10913 (see Chapter 2). The reading frame for the cDNA encoding DmNoppl40-RGG stretches from 124 bp-2187 bp (gray box). Primers used to sequence the cDNA are shown (small boxes). Direction of sequencing is indicated by arrows on the top of primer boxes. The central region between the primers 72ACONNECT and 3'NOPPA and the distal 5' and 3' ends were determined from sequencing reactions using the T3 and T7 primers.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Location (bp)</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>72ACONNECT</td>
<td>819-837</td>
<td>5'-GGCGGCTGGCTGCTCATCGG-3'</td>
</tr>
<tr>
<td>72AEXT</td>
<td>748-768</td>
<td>5'-GGCTACAGGCTTTGAGCTGCTGG-3'</td>
</tr>
<tr>
<td>5'ENDNOPP</td>
<td>598-618</td>
<td>5'-GGACCTCAACCTTCTTGCGGG-3'</td>
</tr>
<tr>
<td>5'FINISH</td>
<td>441-460</td>
<td>5'-CCTCCTCTGAATCGCTGTCC-3'</td>
</tr>
<tr>
<td>5'ENDEXT</td>
<td>375-394</td>
<td>5'-GTGCAGGAGCTTGCTGCTG-3'</td>
</tr>
<tr>
<td>3'NOPPA</td>
<td>1678-1698</td>
<td>5'-GAGAAGACCGCTCGAGACGC-3'</td>
</tr>
<tr>
<td>3'NOPPB</td>
<td>1918-1937</td>
<td>5'-GGCGATCGCAAGAGCTTTGG-3'</td>
</tr>
<tr>
<td>3'NOPPC</td>
<td>2286-2305</td>
<td>5'-CGATAGAAAAACGCTGCGG-3'</td>
</tr>
<tr>
<td>3'NOPPD</td>
<td>2514-2533</td>
<td>5'-CCTTACGCAATAGTTCCGCC-3'</td>
</tr>
</tbody>
</table>
Vector map of pBKCMV/EGFP-DmNopp140-RGG used for expression of GFP-tagged DmNopp140-RGG in Schneider II cells or HeLa cells. P<sub>CMVIE</sub>, Human cytomeglia virus immediate early promoter; lac', alpha complementation portion of β-galactosidase gene; EGFP, enhanced green fluorescent protein gene; DmNopp140-RGG, DmNopp140-RGG cDNA from LD10913; lacZ, β-galactosidase gene; SV40 poly A, SV40 early mRNA polyadenylation signal; f1 origin, f1 single-strand DNA origin; SV40 ori, SV40 origin of replication; Kan<sup>+</sup>/Neo<sup>+</sup>, kanamycin/neomycin resistance gene; TK polyA, thymidine kinase polyadenylation signal.
Vector map of pDsRed1-C1/DmNopp140-RGG used for expression of a DsRed tagged version of DmNopp140-RGG in Schneider II cells or HeLa cells. 

- **PCMVIE**: Human cytomeglo virus immediate early promoter; 
- **DsRed**: red fluorescent protein gene; 
- **DmNopp140-RGG**: DmNopp140-RGG cDNA from LD10913; 
- **SV40 poly A**: SV40 early mRNA polyadenylation signal; 
- **fl origin**: fl single-strand DNA origin; 
- **SV40 ori**: SV40 origin of replication; 
- **Kanr/Neo**: kanamycin/neomycin resistance gene; 
- **HSV TK poly A**: Herpes simplex virus thymidine kinase polyadenylation signal; 
- **pUC ori**: pUC plasmid replication origin.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Vector map of pET30/DmNopp140-RGG used for expression of His-tagged DmNopp140-RGG in *E. coli*. 6 His, 6 histidine tags for purification of *E. coli* expressed DmNopp140 on nickel columns; DmNopp140-RGG, DmNopp140-RGG cDNA from LD10913; fl origin, fl single-strand DNA origin; kan', kanamycin resistance gene; ori, pUC plasmid replication origin; lac I, β-galactosidase repressor.
Vector map of pBluescript/EGFP-DmNopp140-RGG used for synthesis of mRNA encoding a GFP-tagged version of DmNopp140-RGG in \textit{in vitro} run-off transcription reactions. The mRNAs were subsequently injected into \textit{Xenopus} oocytes. The vector should be linearized with \textit{Eco} RI for transcription reactions using the T7 promoter. f1 origin, f1 single-strand DNA origin; \(\beta\)-galactosidase gene; EGFP, enhanced green fluorescent protein gene; DmNopp140-RGG, DmNopp140-RGG cDNA from LD10913; ColE1 origin, plasmid replication origin; Amp\(^r\), \(\beta\)-lactamase gene.
Vector map of pEGFP/DmNopp140 used for expression of a GFP-tagged version of DmNopp140 in Schneider II or HeLa cells. P<sub>CMVIE</sub>, Human cytomeglia virus immediate early promoter; EGFP, enhanced green fluorescent protein gene; DmNopp140, DmNopp140 cDNA from SD10348; SV40 poly A, SV40 early mRNA polyadenylation signal; fl origin, fl single-strand DNA origin; SV40 ori, SV40 origin of replication; Kan<sup>r</sup>/Neo<sup>r</sup>, kanamycin/neomycin resistance gene; HSV TK poly A, Herpes simplex virus thymidine kinase polyadenylation signal; pUC ori, pUC plasmid replication origin
Vector map of pDsRed1-C1/DmNopp140 used for expression of an RFP-tagged version of DmNopp140 in Schneider II cells or HeLa cells. P<sub>CMVIE</sub>, Human cytomegalia virus immediate early promoter; DsRed1, red fluorescent protein gene; DmNopp140, DmNopp140 cDNA from SD10348; SV40 poly A, SV40 early mRNA polyadenylation signal; f1 origin, f1 single-strand DNA origin; SV40 ori, SV40 origin of replication; Kan<sup>r</sup>/Neo<sup>r</sup>, kanamycin/neomycin resistance gene; HSV TK poly A, Herpes simplex virus thymidine kinase polyadenylation signal; pUC ori, pUC plasmid replication origin.
Vector map of pBluescript/EGFP-DmNopp140 used for synthesis of mRNA encoding a GFP-tagged version of DmNopp140 in \textit{in vitro} run-off transcription reactions. The mRNAs were subsequently injected into \textit{Xenopus} oocytes. The vector should be linearized with \textit{Xho} I for transcription reactions using the T7 promoter. \textit{fl} origin, \textit{fl} single-strand DNA origin; \textbeta-galactosidase gene; EGFP, enhanced green fluorescent protein gene; DmNopp140, DmNopp140 cDNA from SD10348; ColE1 origin, plasmid replication origin; Amp\textsuperscript{r}, \textbeta-lactamase gene.
Vector map of pEGFP/DmNopp140 ΔRGG used for expression of a GFP-tagged version of DmNopp140 ΔRGG in Schneider II or HeLa cells. P<sub>CMVIE</sub>, Human cytomeglia virus immediate early promoter; EGFP, enhanced green fluorescent protein gene; DmNopp140 delta RGG, DmNopp140 ΔRGG cDNA from PCR to truncate the RGG domain from DmNopp140-RGG (LD10913); SV40 poly A, SV40 early mRNA polyadenylation signal; f1 origin, f1 single-strand DNA origin; SV40 ori, SV40 origin of replication; Kan<sup>r</sup>/Neo<sup>r</sup>, kanamycin/neomycin resistance gene; HSV TK poly A, Herpes simplex virus thymidine kinase polyadenylation signal; pUC ori, pUC plasmid replication origin.
Vector map of pEGFP/Xenopus coilin used for expression of a GFP-tagged version of Xenopus coilin in Schneider II or HeLa cells. **PCMVIE**, Human cytomegalia virus immediate early promoter; **EGFP**, enhanced green fluorescent protein gene; **Xenopus coilin**, Xenopus coilin cDNA from Drs. Joe Gall and Zheng'an Wu; **SV40 polyA**, SV40 early mRNA polyadenylation signal; **fl origin**, fl single-strand DNA origin; **SV40 ori**, SV40 origin of replication; **Kanr/Neo**, kanamycin/neomycin resistance gene; **HSV TK polyA**, Herpes simplex virus thymidine kinase polyadenylation signal; **pUC ori**, pUC plasmid replication origin.
Vector map of pDsRed1-C1/Xenopus coilin used for expression of a RFP-tagged version of *Xenopus* coilin in Schneider II or HeLa cells. P<sub>CMVIE</sub>, Human cytomegalia virus immediate early promoter; DsRed1, red fluorescent protein gene; *Xenopus* coilin, *Xenopus* coilin cDNA from Drs. Joe Gall and Zheng' an Wu; SV40 poly A, SV40 early mRNA polyadenylation signal; f1 origin, f1 single-strand DNA origin; SV40 ori, SV40 origin of replication; Kan'/Neo', kanamycin/neomycin resistance gene; HSV TK poly A, Herpes simplex virus thymidine kinase polyadenylation signal; pUC ori, pUC plasmid replication origin.
Vita

John Mark Waggener attended Louisiana State University in Baton Rouge, Louisiana, and graduated in December, 1994, with a bachelor's of science in microbiology. He worked in the environmental field before he returned to graduate school at Louisiana State University in Baton Rouge, Louisiana, in August of 1996. Mr. Waggener will graduate with the degree of Doctor of Philosophy in biochemistry from Louisiana State University in August, 2001.
Candidate: John Mark Waggener

Major Field: Biochemistry

Title of Dissertation: Two Splice Variants of Nopp140 in Drosophila melanogaster

Approved:

[Signatures]

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

June 28, 2001