Two splice variants of Nopp140 in *Drosophila melanogaster*

John M. Waggener and Patrick J. DiMario*

Department of Biological Sciences  
202 Life Sciences Building  
Louisiana State University  
Baton Rouge, LA 70803-1715

*To whom correspondence should be addressed.  
Tel: 225-578-1512  
Fax: 225-578-2597
ABSTRACT

The Nopp140 gene of Drosophila maps within 79A5 of chromosome 3. Alternative splicing yields two variants. DmNopp140 (654 residues) is the sequence homologue of vertebrate Nopp140. Its carboxy terminus is 64% identical to that of the prototypical rat Nopp140. DmNopp140-RGG (688 residues) is identical to DmNopp140 throughout its first 551 residues, but its carboxy terminus contains a glycine/arginine rich domain that is often found in RNA-binding proteins such as vertebrate nucleolin. Both Drosophila variants localize to nucleoli in Drosophila Schneider II cells and Xenopus oocytes, specifically within the DFCs. In HeLa cells, DmNopp140-RGG localizes to intact nucleoli, while DmNopp140 partitions HeLa 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. When co-expressed, both Drosophila variants co-localize to HeLa cell nucleoli. Both variants fail to localize to endogenous Cajal bodies (CBs) in Xenopus oocyte nuclei and in HeLa cell nuclei. Endogenous HeLa coilin, however, accumulates around the periphery of phase-light regions in cells expressing DmNopp140. The carboxy truncation (DmNopp140ΔRGG) also fails to localize to CBs, but it forms similar phase-light regions that peripherally accumulate endogenous coilin. Conversely, we see no unusual accumulation of coilin in cells expressing DmNopp140-RGG.
INTRODUCTION

Our traditional understanding of nucleolar function has been the multistage biosynthesis of ribosomes (reviewed by Busch and Smetana, 1970; Hadjiolov, 1985). Ribosomal RNA transcription occurs on the boundaries between the fibrillar centers (FCs) and the dense fibrillar components (DFCs) (Dundr and Raska, 1993; Hozák et al., 1994; Shaw and Jordan, 1995). Site-specific ribose methylation (reviewed by Smith and Steitz, 1997; Weinstein and Steitz, 1999) and pseudouridine conversion (reviewed by Bachellerie and Cavaillé, 1997) are two post-transcriptional modifications of the pre-rRNA that occur within the DFCs. Cleavage of the pre-rRNA, also within the DFCs, yields mature 18S, 5.8S, and 28S rRNAs (e.g. Kass et al., 1990; Mougey et al., 1993; Peculis and Steitz, 1993). Approximately 80 ribosomal proteins assemble with these rRNAs and the extranucleolar-expressed 5S rRNA to form the small (18S rRNA) and large (5S, 5.8S, and 28S) ribosomal subunits. Immature subunits appear in the peripheral granular components (GCs) of nucleoli prior to their further maturation and export to the cytoplasm. Besides ribosome biosynthesis, however, novel functions have been ascribed to nucleoli. They include nuclear import and export, gene silencing, assembly of signal recognition particles, modifications to small non-nucleolar RNAs, cell cycle regulation, and aging (reviewed by Pederson, 1998; Garcia and Pillus, 1999; Johnson et al., 1999; Pederson and Politz, 2000; Visintin and Amon, 2000). Since the early 1970’s (i.e. Orrick et al., 1973), several non-ribosomal nucleolar proteins have been described in yeasts and metazoans (reviewed by Olson, 1990; Olson et al., 2000). How these various proteins participate in the diversified tasks of the nucleolus remains the focus of intense investigation.

Nucleolin, fibrillarin, B23, and Nopp140 are the most extensively studied nucleolar proteins in vertebrates. Nucleolin (110 kDa/pI 5.5; its homologues are Nsr1 in S. cerevisiae and gar2 in S. pombe) is modular in composition (Lapeyre et al., 1987). Its amino terminal third contains alternating acidic and basic domains, its central domain contains four consensus RNA-binding domains (RBDs, see Burd and Dreyfuss, 1994), and its carboxy terminus is rich in glycine and dimethylarginine residues that form several Arg-
Gly-Gly (RGG) motifs. Such RGG motifs are common to a variety of RNA-binding proteins (Burd and Dreyfuss, 1994). Nucleolin interacts with nascent pre-rRNA (Herrera and Olson, 1986; Ghisolfi-Nieto et al., 1996) to facilitate early site-specific cleavages and perhaps other processing events (Ginisty et al., 1998; reviewed by Ginisty et al., 1999). Fibrillarin (34 kDa/pI 8.5; its homologue is Nop1 in yeast) is intimately associated with box C/D snoRNPs for either cleavage or site-specific methylation of the pre-rRNA. In fact, the fibrillarin homologue in the hyperthermophile, *Methanococcus jannaschii*, may be the methyltransferase itself (Wang et al., 2000). Like nucleolin, fibrillarin contains RGG motifs, but within its amino terminus instead of its carboxy terminus. B23 (38 kDa/pI 5.1) is a putative ribosome assembly factor (reviewed by Olson et al., 2000). It contains two acidic regions, but no RBDs or RGG motifs. B23 nevertheless binds nucleic acids (Dumbar et al., 1989; Wang et al., 1994) and displays nuclease activity (Herrera et al., 1995; Savkur and Olson, 1998) as well as chaperone functions (Szepesi and Olson, 1999). Interestingly, B23 may also participate in centrosome duplication in early to mid-G1 (Zatsepina et al., 1999; Okuda et al., 2000).

Meier and Blobel (1990; 1992) described the prototypical Nopp140 (nucleolar phosphoprotein of 140 kDa) in rat. More recently, vertebrate homologues have been identified in *Xenopus* (xNopp180; Cairns and McStay, 1995) and human, (p130; Pai et al., 1995; Chen et al., 1999). Other vertebrate proteins similar and perhaps identical to Nopp140 have been described (Pfeifle et al., 1984, 1986; Vandelaer and Thiry, 1998; Isaac et al., 2000). Nopp140 contains a large central region consisting of several (10-18) alternating acidic and basic regions. The acidic regions contain exclusively aspartic acid, glutamic acid, and serine. The serines are phosphorylated in vivo by casein kinase type II (CKII) enzymes (Meier, 1996; Li et al., 1997), and the resulting phosphoserine residues lend to the acidic property of the region. The interspersed basic regions are rich in lysine, alanine, and proline. A conserved carboxy terminus follows the central acidic and basic domain (Meier, 1996). A putative protein kinase C phosphorylation site is conserved within the carboxy domain, suggesting that Nopp140 is a terminal substrate in signal transduction phosphorylation cascades (Meier, 1996). Like B23, Nopp140 contains no
consensus RBDs or RGG motifs. Unlike B23, however, Nopp140 does not appear to be an RNA-binding protein. Srp40 (41 kDa) in *S. cerevisiae* is the immunological and structural homologue of mammalian Nopp140 (Meier, 1996). Srp40 consists of two relatively long acidic regions that alternate with two short basic regions. The carboxy terminal region of Srp40 is 59% identical to the prototypical terminus in rat Nopp140. Deletion of the *SRP40* gene causes minor growth impairment, whereas over-production of Srp40 causes severe growth impairment (Meier, 1996).

The precise functions of Nopp140 remain uncertain, and our best understanding regarding its function derives from its associations with other nuclear and nucleolar proteins. First, Nopp140 localizes to nucleolar DFCs (Meier and Blobel, 1992). Reports indicate that Nopp140 interacts with the largest subunit of RNA polymerase I (RPA194) (Chen et al., 1999). It may also interact with C/EBPβ and TFIIB to activate the alpha-1-acid glycoprotein gene (*agp*) in mammalian liver (Miau et al., 1997). Based upon these observations, Nopp140 may function in transcription regulation of ribosomal RNA genes within nucleoli and certain non-ribosomal genes presumably outside the nucleolus.

Alternatively, Nopp140 may function in pre-rRNA processing and ribosome biogenesis within the DFCs. Nopp140 associates with both classes of mammalian snoRNP particles (box H/ACA and box C/D snoRNPs) as determined by co-immunoprecipitations (Yang et al., 2000). For example, Nopp140 associates in stoichiometric amounts with rat NAP57 (*Nopp140*-associated protein of 57 kDa) (Meier and Blobel, 1994), a protein component of box H/ACA snoRNPs. Interestingly, rat NAP57 shares conserved regions with the *S. cerevisiae* Cbf5 gene product that may play several roles in the transcription, processing, and pseudouridylation of yeast pre-rRNA (Cadwell et al., 1997; Lafontaine et al., 1998). The human homologue of NAP57, called dyskerin, is a product of the *DKC1* gene. Mutations in *DKC1* lead to dyskeratosis congenita, a rare X-linked (Xq28) recessive disease in which progressive bone-marrow failure is the primary cause of mortality. The *Drosophila* homologue of NAP57 is Nop60B (Phillips et al., 1998), the product of the *minifly* (*mfl*) gene that maps to the right arm of chromosome 2 within region 60B-C (thus
its designation, Nop60B) (Giordano et al., 1999). Mutations in *mfl* lead to reduced body size, abnormal eggs, and reduced fertility. Interestingly, Cbf5, NAP57, dyskerin, and Nop60B are all related to TruB, a pseudouridine synthase for tRNAs in *E. coli*. The implication here is that Nopp140 may be involved with pseudouridine conversion by the box H/ACA snoRNP particles. The yeast homologue, Srp40, also associates with box H/ACA snoRNP particles (Yang et al., 2000). Besides box H/ACA snoRNPs, Nopp140 co-precipitates with components of the box C/D snoRNPs. Specifically, Nopp140 associates with fibrillarin and the newly characterized mammalian NAP65 (Nop5/58p in yeast), both of which are components of box C/D snoRNPs (Yang et al., 2000).

In addition to these associations, vertebrate Nopp140 shuttles between the nucleus and the cytoplasm (Meier and Blobel, 1992). This observation is consistent with a chaperone function of Nopp140 in the transport of karyophilic proteins into the nucleus, or in the export of nuclear products to the cytoplasm. Phosphorylation may regulate this chaperone function. That is, Nopp140 binds NLS-containing proteins in vitro when its multiple serines are phosphorylated by CKII, while the dephosphorylated version of Nopp140 fails to do so (Meier and Blobel, 1990).

One of the most intriguing interactions exists between Nopp140 and p80 coilin (Isaac et al., 1998). Coilin is a marker protein for nuclear organelles that were traditionally called coiled bodies (Andrade et al., 1991), but recently renamed Cajal bodies (Gall et al., 1999) in honor of S.R. Ramón y Cajal who first described the structures as accessory organelles to the nucleoli (Ramón y Cajal, 1903). Cajal bodies (CBs) are often found in the vicinity of nucleoli, adjoining nucleoli, or within the nucleoli of plant and animal cells (reviewed by Matera, 1999; Gall, 2000). The molecular composition of CBs continues to unfold (Bohmann et al., 1995a; Gall et al., 1999; Matera, 1999; Gall, 2000). Besides the aforementioned p80 coilin, somatic cell CBs contain pre-mRNA splicing components such as snRNAs U1, U2, U4, U5, U6, the trimethylguanosine cap epitope, and Sm proteins (Matera, 1999; Gall, 2000). CBs also contain nucleolar proteins Nopp140, NAP57, fibrillarin, Gar1p, topoisomerase I, and the ribosomal protein, S6. CBs contain snoRNAs
U3 and U8 (Narayanan et al., 1999; Speckman et al., 1999), but somatic cell CBs do not contain nucleolin, B23, or any ribosomal RNA. CBs, therefore, are not directly involved in ribosome biogenesis. Interactions between CBs and nucleoli are well established (Bohmann et al., 1995b; Sleeman et al., 1998; Isaac et al., 1998; Platani et al., 2000). Deletion mutants of p80 coilin altered CBs and nucleoli (Bohmann et al., 1995b; Sleeman et al., 1998). Conversely, deletion mutants of Nopp140 caused dominant negative effects on the normal distribution of nucleolar proteins within nucleoli and CBs (Isaac et al., 1998). Physical interactions were observed by Platani et al. (2000) who used time-lapse fluorescence microscopy to show that CBs move to and from nucleoli. Association of Nopp140 with both nucleoli and CBs supports the hypothesis that Nopp140 shuttles RNA processing complexes (snoRNPs) to and from nucleoli.

The conclusion from these introductory comments is that Nopp140 appears to have multiple and diverse functions. Here we introduce two splice variants of Nopp140 in Drosophila melanogaster that differ in their carboxy ends. DmNopp140 appears to be the sequence homologue of vertebrate Nopp140 in overall peptide domain composition and arrangement. DmNopp140-RGG is identical to DmNopp140 throughout most of its primary sequence (residues 1-551), but its carboxy terminal tail contains a RGG domain that is highly reminiscent of the carboxy RGG domain in vertebrate nucleolin (Lapeyre et al., 1987). As far as we know, this is the first example of a Nopp140-like protein that contains a peptide domain typically reserved for RNA-binding proteins (Burd and Dreyfuss, 1994). Comparative molecular and genetic analyses of the two Drosophila Nopp140 variants should provide valuable insights to Nopp140’s diverse functions, while at the same time expanding our knowledge of nucleolar functions, both traditional and novel.
**MATERIALS AND METHODS**

*Recovery, Sequencing, and Cloning of Drosophila Nopp140 cDNAs.*

We used standard molecular biology techniques (Ausubel et al., 1987-1995) to screen aliquots of a *Drosophila melanogaster* stage 10 egg chamber cDNA lambda phage library. The probe was a random primed, $[^{32}\text{P}]$-labeled subclone of our *Xenopus* nucleolin cDNA (Rankin et al., 1993; accession number X63091). Specifically, the subclone was a 444 bp fragment that spanned the *Nco*I site at the translation start site to a downstream *Pst*I site. It encodes most of the alternating acidic and basic regions within the amino terminal third of the smaller of two *Xenopus* nucleolin proteins (see Meßmer and Dreyer, 1993 for a comparison of the two nucleolin proteins in *X. laevis*). We used low stringency washes (2 X SSC without SDS at room temperature) to detect related *Drosophila* cDNA sequences. Four strongly positive plaques were picked and re-screened, again under low stringency to establish clonal purity. Individual plaques were amplified, and phage DNA was prepared and digested with *Eco*RI to liberate the *Drosophila* cDNAs from the lambda genome. The *Drosophila* inserts were ligated into pBluescript KS(+) (Strategene, La Jolla, CA) and sequenced in both directions using Sanger’s dideoxy method for DNA sequencing. We used Sequenase (USB, Cleveland, OH) according to the manufacturer’s recommendations.

One of the *Drosophila* inserts that displayed a strong hybridization signal with the *Xenopus* probe was only 787 bp in length (B72A). Its deduced translation product contained alternating acidic and basic regions, and thus it was highly reminiscent of the alternating acidic and basic regions within vertebrate nucleolin and Nopp140. We used this insert to rescreen the *Drosophila* cDNA library, this time using higher stringency washes (0.5 X SSC, 0.1 X SDS at 60° C). Rescreening identified several larger inserts that we sequenced. One of the inserts provided a nearly full-length cDNA that encoded a Nopp140-like protein. The deduced protein sequence, however, contained a RGG carboxy terminus, and we refer to the protein as DmNopp140-RGG. To provide the
missing 5’ end of the cDNA, we obtained an expressed sequence tag (LD10913) from Genome Systems Inc., (St. Louis, MO) that proved to be a complete cDNA encoding DmNopp140-RGG (our accession number AF162774).

While sequencing the library's cDNA that encodes DmNopp140-RGG, the Berkeley Drosophila Genome Project (BDGP) published the Drosophila genome (Flybase, 1999). We used the cDNA sequence for DmNopp140-RGG in a BLAST search of the genome, and found the Nopp140 gene in polytene region 79A5 on the proximal left arm of chromosome 3. The BDGP predicted two conceptual transcripts from this gene. They are splice variants that encode DmNopp140-RGG (CT22845) and DmNopp140 (CT22833). We next obtained several additional ESTs from Research Genetics (Huntsville, AL), and used restriction enzyme digestion patterns to determine which of the several clones expressed the two proteins (see Fig. 1B).

Full-length cDNAs encoding DmNopp140 (EST SD10348) and DmNopp140-RGG (EST LD10913) were ligated into both pEGFP-C3 and pDsRed1-C1 (Clontech, Palo Alto, CA). This allowed us to express the red shifted version of the green fluorescent protein (EGFP) or the red fluorescent protein (RFP) fused in frame to the amino terminal ends of DmNopp140 and DmNopp140-RGG. In order to engineer a carboxy terminal truncation of DmNopp140-RGG (referred to as DmNopp140ΔRGG), two PCR primers were designed to amplify the DmNopp140-RGG cDNA (LD10913) except for the sequences that encode the RGG tail. The upstream primer was complementary to the non-coding strand, and it contained an EcoRI site (underlined) just upstream of the ATG start codon (in bold) (5’-GCGAATTCTCATGACAGACCTGCTAAAGATAGCC-3’). The downstream primer (5’-AAGGATCCTTATCCGTTGTTGTGCTTCTTAAAGTCG-3’) was complementary to the coding strand, and it contained a stop codon (in bold) that would have normally encoded amino acid residue 562 of DmNopp140-RGG. The underlined BamHI site was included for cloning purposes. The resulting PCR product was digested with EcoRI and BamHI and then ligated into the pEGFP-C3 vector at the respective restriction sites. The EcoRI site in the upstream PCR primer was positioned such that the
ATG start codon was in frame with the sequences that encode the EGFP. The region of DmNopp140-RGG that is deleted in DmNopp140ΔRGG is shown as italicized letters in Fig. 1C. All recombinant plasmids were purified twice by cesium banding prior to their use in transfection assays (see below). The CMV immediate early promoter within pEGFP-C3 and pDsRed1-C1 controls expression of the fusion proteins even in Drosophila cells (see Echalier, 1997).

Bacterial Expression and In Vitro Phosphorylation

The full-length cDNA encoding DmNopp140-RGG was removed from EST LD10913 using a BspHI site that is located at the translation start site and a BamHI site within the 3' UTR. The cDNA was ligated into pET30 (Novagen, Madison, WI) at the NcoI (compatible with BspHI) and BamHI sites such that the DmNopp140-RGG coding sequence was positioned in frame with sequences within pET30 that encode the His tag. We transformed E. coli strain JM109 DE3 pLysS with the recombinant pET30 plasmid. To express DmNopp140-RGG, 5 ml of an overnight culture was transferred to 1 liter of LB broth. The culture was allowed to grow to an OD\textsubscript{600} of 0.6 at which point IPTG was added to a final concentration of 1 mM. Cells were incubated at 37\degree C for an additional 3 hrs and then harvested by centrifugation at approximately 6000 X g. Cells 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 \mu g/ml pepstatin, and 1.0 \mu g/ml leupeptin. Cells were lysed by sonication for 2 min on ice. The cell lysate was cleared by centrifugation at 25,000 X g for 20 min at 4\degree C. The supernatant was passed through a 0.4 \mu m syringe filter directly onto a HIS-tag column using Novagen’s recommendations. Eluate was collected and dialyzed overnight at 4\degree C in 75 mM KCl, 10 mM Tris HCl, pH 7.2, 1 mM EDTA, and 0.1 mM PMSF. EDTA was included in the dialysis buffer to chelate nickel that leaches off the column. We concentrated the protein by placing the dialysis bag (molecular weight cut-off of 13,000 to 14,000) into solid polyvinylpyrrolidone
(PVP, average molecular weight = 360,000). This resulted in rapid dehydration without adversely changing ionic strength.

The enriched protein was phosphorylated in vitro using gamma labeled $[^{32}P]$-ATP at 800 Ci/mmol (ICN Pharmaceuticals, Costa Mesa, CA) and either casein kinase II or Cdk1/cyclin B protein kinase (MPF). Both enzymes were purchased from New England Biolabs, Beverly, MA. Phosphorylations were performed according to the manufacturer’s recommendations.

**Cell Culture and Transient Transfection**

All media and antibiotics were from Life Technologies GibcoBRL, Gaithersburg, MD. *Drosophila* Schneider II cells were grown in Schneider’s *Drosophila* medium supplemented with 10% fetal calf serum (FCS) and 50 µg/ml penicillin-streptomycin-glutamine in a 25°C ambient air incubator. HeLa cell stocks were maintained at 37°C in 5% CO$_2$ using Delbecco's Modified Eagles Medium (DMEM) that was supplemented with 10% FCS and 50 µg/ml gentamycin.

For transfection of HeLa cells, approximately 1 X 10$^5$ cells were grown on 22 X 22 mm coverslips in six well culture plates (e.g. Falcon 3046) at 37°C in a 5% CO$_2$ atmosphere. Transfection was by DNA-calcium phosphate precipitation using the N,N-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES) method of Chen and Okayama (1988). After adding the DNA precipitate, cells were incubated at 35°C in 3% CO$_2$ for 12 to 18 hours (overnight). Cells were then washed twice with 1 X PBS and either fixed with 2.0% paraformaldehyde in 0.6 X PBS for 1.5 h at room temperature, or re-cultured in DMEM with 10% FCS at 37°C in a 5% CO$_2$ atmosphere for an additional 24 h prior to 1 X PBS washing and 2% paraformaldehyde fixation. Transfection methods used for mammalian cells were used for the Schneider II cells, except that the *Drosophila* cells were
maintained in Schneider’s *Drosophila* medium and in ambient air at 25°C throughout all procedures.

*Antibodies and Antibody Staining*

Following formaldehyde fixation, cells (still attached to cover slips) were washed several times with 1 X PBS and then treated with 0.1% Triton X100 in 1 X PBS for 5 min. The cells were again washed with 1 X PBS and blocked with 10% horse serum or 3% BSA in 1 X PBS. Anti-human nucleolin was a gift from Dr. Benigno Valdez, Baylor College of Medicine, Houston, TX. The S4 anti-human fibrillarin (Ochs et al., 1985; Lischwe et al., 1985) was a gift from Dr. Robert Ochs, Scripts Research Institute, La Jolla, CA. The anti-human p80 coilin rabbit serum (R288) was developed by Andrade et al. (1993), but provided to us by Dr. Joe Gall, Carnegie Institution of Washington, Baltimore, MD. Primary antibodies were diluted appropriately in 10% horse serum or 3% BSA and placed onto the coverslips for 1-2 h at room temperature. Respective affinity purified secondary antibodies (ICN/Cappel, Costa Mesa, CA) were tagged either with fluorescein or rhodamine.

*In Vitro mRNA Production and Oocyte Injection*

We prepared synthetic mRNAs encoding either the GFP-tagged DmNopp140 or the GFP-tagged DmNopp140-RGG using a T7 Message Machine kit from Ambion, Inc. (Austin, TX). DNA inserts encoding the GFP tag along with either DmNopp140-RGG or DmNopp140 were ligated into pBluescript (Stratagene, La Jolla, CA) to make use of its T7 promoter for in vitro transcription as described by Ambion. We injected the transcripts into *Xenopus* oocytes using procedures described by Heine et al. (1993). After injection the oocytes were cultured overnight at 18°C in OR2 medium. The next day, oocyte nuclear contents were prepared for light microscopy as described by Gall (1998).
RESULTS

Recovery and Sequence of DmNopp140-RGG

In an effort to recover *Drosophila* cDNAs that encode either nucleolin or Nopp140, we originally used a 5’ subclone of our *Xenopus* nucleolin cDNA (Rankin et al., 1993) to screen a stage 10 egg chamber cDNA library under low stringency. The *Xenopus* subclone encodes the majority of the alternating acidic and basic regions within the amino terminal third of *Xenopus* nucleolin. We recovered and sequenced several overlapping *Drosophila* inserts. Our full length cDNA (acc. no. AF162774) encodes what we refer to as DmNopp140-RGG. The *Drosophila* protein resembles vertebrate Nopp140 (Meier, 1996) in that it contains a long central series of alternating acidic and basic regions (Fig. 1A).

We used the DmNopp140-RGG cDNA sequence in a BLAST search of the Berkeley *Drosophila* Genome Project (BDGP). Conceptual gene CG7421 maps within polytene region 79A5 on the left arm of polytene chromosome 3 (see the various links at http://hedgehog.lbl.gov:8000/cgi-bin/annot/gene?CG7421). The pre-mRNA is 3247 nts with the translation start codon positioned at residues 88-90. The BDGP lists two splice variants for the gene, conceptual transcripts CT22833 and CT22845. CT22845 encodes DmNopp140-RGG, and its sequence is nearly identical to our cDNA sequence. Three exons constitute CT22845; they include nucleotides 88-186, 389-1942, and 2580-2993 of the pre-mRNA. These exons encode amino acids 1-33, 34–551, and 552–688, respectively (see Fig. 1C).

Conversely, CT22833 encodes DmNopp140 (Fig. 1A and D) with a carboxy terminal tail that is highly conserved among previously identified Nopp140 proteins (Meier, 1996; see Table I). Four exons constitute CT22833. The first two exons include
nucleotides 88-186, and 389-1942, and again they encode amino acid residues 1-33 and 34-551 as in CT22845. The next two exons include nucleotides 2184-2342 and 3096-3248, and they respectively encode amino acids 552-604 and 605-654 within DmNopp140 (see Fig. 1D). Thus, the two splice variant mRNAs, CT22845 and CT22833, are identical in their first two exons that encode amino acid residues 1-551. Due to alternative splicing, however, the two transcripts contain mutually exclusive exons that encode totally different carboxy termini. The two Drosophila proteins differ in sequence beginning at residue 552. The dark gray boxes in Figs. 1C and D contain residue 551 that is common to both proteins and residue 552 that differs between the two proteins.

Two lines of evidence from the BDGP indicate that both mRNAs are expressed. First, Genscan or Genie programs used by the BDGP suggest that CG7421 encodes both transcripts. Second, expressed sequence tags (ESTs) exist for both transcripts. We purchased several ESTs from Genome Systems Inc., (St. Louis, MO) and Research Genetics (Huntsville, AL), and used the predicted restriction enzyme digestion patterns (CT22833 versus CT22845) to verify that at least one of the purchased clones encoded DmNopp140 versus DmNopp140-RGG (Fig. 1B).

DmNopp140-RGG contains 688 amino acid residues while DmNopp140 is 654 residues in length (Fig. 1C and D). The prototypical rat Nopp140 is longer with 704 residues. Both DmNopp140-RGG and DmNopp140 contain alternating acidic (light gray shading) and basic regions that constitute a large central domain in both proteins. The acidic regions within the two Drosophila proteins are similar to those within rat Nopp140; they are rich in glutamic acid, aspartic acid, and serine. Also like rat Nopp140, the basic regions in the two Drosophila variants are rich in alanine, proline and lysine.

Two readily apparent differences exist between the rat protein and the two Drosophila proteins. The first is the number of alternating acidic and basic regions is greater in the Drosophila proteins (15 acidic regions) versus the rat protein (12 acidic regions). We note, however, that Xenopus xNopp180 has 18 acidic regions (Cairns and
McStay, 1995). The second difference is that the overall length of the central acidic-basic domain in the two Drosophila variants is shorter when compared to the central domains in the rat and Xenopus proteins. Although the differences may seem minor, they could explain some unexpected results pertaining to the failure of these proteins to localize to CBs in Xenopus oocytes and HeLa cells (see below).

The carboxy terminus of DmNopp140 is fairly well conserved across the eukaryotes (Meier, 1996; Isaac, et al., 1998). FASTA results display 65 and 64 percent identity over 94 amino acid stretches when the carboxy tail of DmNopp140 (beginning at residue 552) is compared to that of human and rat Nopp140 proteins, respectively (Table I). Further, Meier (1996) and Isaac et al. (1998), described two separate sub-domains that comprise the carboxy terminus of Nopp140 homologues. The first half of the carboxy terminus, referred to as C$_a$ (Isaac et al., 1998) shows good homology among vertebrates (e.g. 54% identity between rat and Xenopus). The latter half of the carboxy domain, referred to as C$_b$, has good homology throughout the eukaryotes in general (e.g. 59% identity between rat and yeast) with even higher degrees of homology between the metazoans (e.g. 81% identity between rat and Xenopus) (Meier, 1996). Interestingly, the breakpoint between the peptide encoded by the third exon (residues 552-604 = DmNopp140-C$_a$) and the peptide encoded by the fourth exon (residues 605-654 = DmNopp140-C$_b$) correlate well with the first and second halves of rat Nopp140. That is, DmNopp140-C$_a$ has 50% identity with the comparable sequence in both human and rat Nopp140, while DmNopp140-C$_b$ has 78% and 76% identities with the comparable sequences in human and rat Nopp140s, respectively. These homology values for DmNopp140-C$_a$ (50%) and DmNopp140-C$_b$ (78%) are in good agreement with the reported metazoan homologies of 54% identity and 81% identity, respectively.
DmNopp140-RGG, on the other hand, contains a glycine and arginine rich (RGG) domain near its carboxy terminus. Although two adjacent RGG motifs reside at position 562-567, most of the RGG motifs in DmNopp140-RGG lie within a defined domain spanning residues 612-669. We call this latter region the RGG domain. The RGG domain in DmNopp140-RGG contains 58 residues; 44 glycines, 10 arginines, and 4 phenylalanines. All 10 arginines within the *Drosophila* RGG domain are followed by at least two glycines. This domain of DmNopp140-RGG is quite similar to the one near the carboxy terminus of vertebrate (CHO) nucleolin (e.g. Lapeyre et al., 1987). The RGG domain in CHO nucleolin is 52 residues in length; it consists of 37 glycines, 10 arginines, and 5 phenylalanines (Lapeyre et al., 1987). Nine of the ten arginines in the CHO RGG domain are followed by at least two glycines. Interestingly, the RGG domain in both DmNopp140 and vertebrate nucleolin do not constitute the very ends of the proteins; 19 residues follow the RGG domain in DmNopp140-RGG, while 12 residues follow the RGG domain in CHO nucleolin. The significance of this terminal carboxy tail remains unknown. We conclude that DmNopp140-RGG is novel; it appears to be a composite of Nopp140 throughout most of its length and a carboxy terminus rich in RGG motifs that are usually reserved for RNA-binding proteins such as nucleolin (Burd and Dreyfuss, 1994).

**MPF and CKII In Vitro Phosphorylations**

The deduced peptide sequence of DmNopp140-RGG indicated eight putative MPF (cdk1/cyclin B) phosphorylation sites within the repeating basic regions (bold SPKK or TPAK sites in Fig. 1C). Similar sites have been mapped within the basic regions of vertebrate nucleolin (Belenguer et al., 1990; Peter et al., 1990; Zhu et al., 1999). To show that DmNopp140-RGG is an in vitro substrate for MPF, we expressed DmNopp140-RGG as an amino terminal His-tagged fusion protein in *E. coli*, and then purified the protein using a nickel affinity column. The purified protein was labeled in vitro by MPF in the presence of gamma-[\(^{32}\)P] ATP (Fig. 2, lanes A and A’). If it occurs in vivo, MPF phosphorylation of Nopp140 may regulate nucleolar disassembly during prophase, while its dephosphorylation may regulate nucleologenesis during telophase (see Pai et al., 1995).
Rat Nopp140 is a substrate for casein kinase type II (CKII) enzymes (Meier, 1996). CKII specifically phosphorylates the multiple serine residues within the central acidic regions. CKII readily phosphorylated the His-tagged DmNopp140-RGG in vitro (Fig. 2, lanes B and B’). The extent of phosphorylation was much greater with CKII than with MPF (compare lanes A’ and B’). Although we can not rule out differences in enzyme activities, the enhanced phosphorylation by CKII versus MPF is consistent with the large number of CKII sites (approximately 82) versus the eight putative MPF sites. Phosphorylation by CKII retarded the mobility of the protein in the SDS-polyacrylamide gel (arrow in lanes B and B’) as previously described for the CKII phosphorylation of rat Nopp140 (Meier, 1996). Both Drosophila variants are identical at these MPF and CKII phosphorylation sites, and we fully expect that DmNopp140 will show the same phosphorylation profile as shown here for DmNopp140-RGG. As with nucleolin (Csermely et al., 1993; Bonnet et al., 1996), CKII phosphorylation of Nopp140 in vivo probably occurs during interphase in response to growth signals.

Exogenous Expression in Drosophila Schneider II Cells

The respective cDNAs encoding either DmNopp140-RGG or DmNopp140 were ligated into expression vectors downstream and in frame with sequences encoding the green fluorescent protein (GFP) or the red fluorescent protein (RFP). We purified the plasmids by cesium chloride gradient ultra-centrifugation and then used them to co-transfect Drosophila Schneider II cells (Fig. 3). Strong upstream CMV promoters directed constitutive transcription for all constructs.
Drosophila Schneider II cells contain one prominent nucleolus. GFP-DmNopp140-RGG (Fig. 3A and B) and GFP-DmNopp140 (Fig. 3C and D) localized to the single nucleolus in the relatively few cells that were transfected. The observation confirmed our expectations that both variants are nucleolar proteins in Drosophila cells. Although the number of transfected Schneider II cells was limiting, we observed no other nuclear structures (e.g. Cajal bodies) in cells exogenously expressing GFP-DmNopp140-RGG or GFP-DmNopp140.

Exogenous Expression in Xenopus Oocytes

Xenopus oocytes contain large multiple extra-chromosomal nucleoli that allow for unencumbered localization of nucleolar components (e.g. Shah et al., 1996). Toward this end, we injected Xenopus oocytes with synthetic transcripts encoding either GFP-DmNopp140 or GFP-DmNopp140-RGG. Following mRNA injection, the oocytes were incubated at 18°C overnight to allow for protein synthesis and cytoplasmic to nuclear translocation, after which the nuclear contents were prepared for light microscopy according to Gall (1998). The nuclear preparations were stained with DAPI to localize the rDNA within the FCs of the multiple nucleoli (Fig. 4C, H). Clearly, GFP-DmNopp140-RGG (Fig. 4B) and GFP-DmNopp140 (Fig. 4G) localized to the DFC regions immediately surrounding the FCs. This is the sub-compartment of the DFC that is enriched for endogenous fibrillarin (see Shah et al., 1996). We know that mammalian Nopp140 associates with fibrillarin in box C/D snRNPs (Yang et al., 2000). Thus the localization of both Drosophila variants to this specific site within the DFCs of Xenopus nucleoli is consistent with the possibility that both Drosophila variants are intimately involved in snoRNP transport and/or in pre-rRNA processing.
In three separate trials, GFP-Nopp140-RGG and GFP-DmNopp140 labeled only nucleoli after 18-24 hours post injection (Fig. 4E and J, respectively). No other germinal vesicle bodies contained either variant. The oocyte CBs (formerly called spheres or C snurposomes, see Gall, 2000) were completely devoid of any GFP labeling (arrows in Fig. 4E and J). This was surprising since the large Cajal bodies in amphibian oocyte nuclei are known to contain Nopp140 (Gall et al., 1999). Isaac et al. (1998) demonstrated a lag time of between 12 and 24 before Nopp140 appeared in the CBs of transfected COS-1 cells, and this might explain our failure to see labeling in the oocyte CBs. On the other hand, Wu et al., (1994) showed a very rapid (1-4 h) localization of coilin to *Xenopus* oocyte CBs after injecting mRNAs. The length of time between our oocyte injections and nuclear preparations (approximately 18 h) should have been sufficient time for at least some of the *Drosophila* proteins to accumulate within the CBs. We know that the central domain of alternating acidic and basic regions of Nopp140 is required for Cajal body localization (Isaac et al., 1998). Perhaps sufficient sequence differences exist between xNopp180 and the two *Drosophila* variants (see above) such that the *Drosophila* proteins fail to localize to CBs in the *Xenopus* oocyte nucleus.

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

We employed the same expression construct used for the Schneider II cells to express and localize DmNopp140-RGG in HeLa cells. As expected, GFP-DmNopp140-RGG localized to the phase-dark nucleoli (Fig. 5A, B). In transfected cells that expressed a moderate amount of GFP-DmNopp140-RGG, we saw a punctate staining pattern over the nucleoli (Fig. 5B) suggesting that GFP-DmNopp140-RGG localized to sub-domains of the nucleoli. These sub-domains may be analogous to the subdivision within the DFCs that we observed in the *Xenopus* oocyte nucleoli (Fig. 4B). Endogenous fibrillarin co-localized well with DmNopp140-RGG in nucleoli of HeLa cells expressing moderate amounts of DmNopp140-RGG (arrow Fig. 5D and E). In HeLa cells expressing greater quantities of GFP-DmNopp140-RGG (arrowhead in Fig. 5D and E), the punctate staining was not
readily apparent due to greater accumulations of GFP-DmNopp140-RGG within the nucleoli. Nevertheless, fibrillarin continued to co-localize.

We counter-stained other transfected HeLa cells with anti-human nucleolin (Fig. 5H). As expected, GFP-DmNopp140-RGG and endogenous nucleolin co-localized to the relatively large, phase-dark nucleoli. Smaller nuclear organelles were evident by phase contrast microscopy (arrows in Fig. 5F), but they did not contain GFP-DmNopp140-RGG (Fig. 5G), nor did they stain with anti-nucleolin (Fig. 5H). This observation is significant because the nucleoplasm in HeLa cells that over-expressed GFP-DmNopp140-RGG was lightly but uniformly labeled (Fig. 5G), and we observed no other nuclear organelle (e.g. CBs) that contained GFP-DmNopp140-RGG other than the large nucleoli that also contained endogenous nucleolin. Since nucleolin is not a component of somatic cell CBs, we conclude that DmNopp140-RGG localized only to nucleoli when exogenously expressed in HeLa cells.

Figure 5

Since DmNopp140-RGG failed to localize to the large CBs in *Xenopus* oocyte nuclei, we critically tested whether or not GFP-DmNopp140-RGG could localize to endogenous CBs in HeLa cells. After transfection, the cells were counter-stained with R288, a rabbit antiserum directed against human p80 coilin (Andrade et al., 1993). Transfected cells again contained nucleoli that were well labeled by GFP-DmNopp140-RGG (Fig. 5J), but the endogenous CBs shown in Fig. 5K were completely devoid of GFP-DmNopp140-RGG. The arrows in Fig. 5I, J, and K are well matched denoting the position of the endogenous CBs. The cells in Fig. 5I-K were washed free of the DNA-Ca$^{2+}$ 24 hours after its addition and then incubated under normal conditions for an additional 24 hours before fixation. This should have been ample time for GFP-DmNopp140-RGG to transit the nucleoli and accumulate within CBs (see Isaac et al. 1998) We observed identical results with transfected cells that were fixed 24 hours after the
DNA-Ca$^{+2}$ was initially added. The failure of GFP-DmNopp140-RGG to localize to endogenous CBs in HeLa cells was consistent with its failure to localize to CBs in *Xenopus* oocytes.

*Exogenous Expression of DmNopp140 in HeLa Cells*

Figure 6

As with GFP-DmNopp140-RGG, we used the same expression construct used for the Schneider II cells to express and localize GFP-DmNopp140 in HeLa cells. Unlike GFP-DmNopp140-RGG, however, GFP-DmNopp140 caused HeLa cell nucleoli to partition into phase-light and phase-dark regions (Fig. 6A and B). GFP-DmNopp140 localized exclusively to the phase-light regions (Fig. 6B). In some cells, the phase-light regions extended well into the nucleoplasm, but they remained attached to the phase-dark regions of the nucleoli (arrows in Fig. 6A, C, F, and I).

We counter stained the partitioned nucleoli with anti-fibrillarin (Fig. 6E) and anti-nucleolin (Fig. 6H) to initially define the molecular composition of the phase-light and phase-dark regions. GFP-DmNopp140 co-localized with fibrillarin in the phase-light regions (compare the matched arrows in Fig. 6D and E). This co-localization is in accord with co-immunoprecipitation results that showed an association between mammalian Nopp140 and fibrillarin within C/D box snoRNP particles (Yang et al., 2000). While GFP-DmNopp140 accumulated within the phase-light regions (Fig. 6G), endogenous nucleolin localized within the phase-dark regions (compare the matched arrows in Fig. 6F-H). The partitioning of endogenous fibrillarin (phase-light region) from endogenous nucleolin (phase-dark region) in HeLa cells expressing GFP-DmNopp140 suggests that nucleolin and fibrillarin do not form tight associations in vivo.
We also tested the ability of GFP-DmNopp140 to localize to endogenous CBs in transfected HeLa cells. Counter-staining the cells with anti-human coilin showed that GFP-DmNopp140, like GFP-DmNopp140-RGG, failed to localize to endogenous CBs (arrowheads in Fig. 6J-K). As with DmNopp140-RGG, its failure to associate with endogenous CBs is consistent with the results obtained with GFP-DmNopp140 expression in *Xenopus* oocytes.

Interestingly, in about half of the transfected cells expressing GFP-DmNopp140, endogenous coilin redistributed to the phase-light regions. There was a mixture of observed coilin re-distributions, and we present these diverse redistributions in right- and left-hand panels for Fig. 6I-K. First, endogenous coilin was observed to accumulate on the very periphery of some phase-light regions. This is evident in the transfected cell in the center of the left-hand panels of Fig. 6I-K. In this particular cell, the nucleolus is partitioned into a phase-dark spot (arrow) and a phase-light(er) region. Anti-coilin labeled only the left side of this phase-light region (same cell in the left-hand panel of Fig. 6K). Conversely, we observed no redistribution of endogenous coilin in the transfected cell in lower portion of the left-hand panels. In this particular cell, the nucleoli were again partitioned, but endogenous coilin remained within small CBs (arrowheads point to 0.5-1 µm spheres in the left-hand panel of Fig. 6K); the anti-coilin completely failed to stain the phase-light regions containing GFP-DmNopp140.

In the right-hand panels of Fig. 6I, J, and K, two transfected cells expressed GFP-DmNopp140, but the partitioning of their nucleoli was more dramatic (arrows point to the phase-dark regions while the phase-light regions spill out into the nucleoplasm). Anti-coilin again stained the periphery of these phase-light regions, but there was now internal staining as well. Occasionally, we observed what appeared to be a Cajal body within the periphery of a phase-light region as seen in the upper-most phase-light region in the right-hand panel of Fig. 6K. In addition to anti-coilin staining the phase-light regions, separate CBs were also evident in cells expressing GFP-DmNopp140 (arrowheads in the lower
nucleus of the right-hand panels). Again, GFP-DmNopp140 failed to localize to these CBs.

In summary, GFP-DmNopp140 causes nucleoli to partition into phase-light and phase-dark regions. GFP-DmNopp140 and fibrillarin co-localize to the phase-light regions while nucleolin localizes to the phase-dark regions. GFP-DmNopp140 fails to localize to endogenous CBs, but endogenous coilin appears to associate with the phase-light regions, preferentially on the periphery of these regions in many cases.

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

We co-expressed GFP-DmNopp140 (Fig. 7B, E, and H) and RFP-DmNopp140-RGG (Fig. 7C, F, and I) in HeLa cells to determine if they would co-localize together within partitioned or intact nucleoli. Both proteins co-localized to fairly intact nucleoli (Fig. 7A) when expressed in approximately equal amounts based on fluorescence intensities (Fig. 7B and C). There were no prominent phase-light nucleolar regions protruding into the nucleoplasm as observed in Fig. 6A. Upon closer examination, however, some nucleoli appeared slightly partitioned, but not much more than what we occasionally observed in non-transfected cells (compare the two arrows in Fig. 7A). Cells that expressed greater amounts of GFP-DmNopp140 (Fig. 7E) relative to RFP-DmNopp140-RGG (Fig. 7F) again contained partitioned nucleoli with prominent phase-light and phase-dark regions (Fig. 7D). Interestingly, DmNopp140-RGG localized with DmNopp140 within these phase-light regions, suggesting an association may occur between these two proteins. Finally, nucleoli appeared morphologically normal in cells that expressed less DmNopp140 (Fig. 7H) as compared to DmNopp140-RGG (Fig. 7I). Both proteins again co-localized to the intact nucleoli in these cells. The observations in Figure 7 collectively indicate that the extent to which nucleoli partition into phase-light and phase-dark regions is...
proportional to the amount of DmNopp140 expressed within the cell, but that an equal co-expression of DmNopp140-RGG may dampen this nucleolar partitioning.

**Localization Patterns of DmNopp140ΔRGG in HeLa Cells**

The carboxy RGG domain of DmNopp140-RGG is very similar to the carboxy RGG domain of vertebrate nucleolin. When the RGG domain of nucleolin is deleted, the resulting truncation translocates to the nucleus (the bipartite NLS is further upstream), but it fails to associate with nucleoli (Meßmer and Dreyer, 1993; Schmidt-Zachmann and Nigg, 1993; Heine et al., 1993; Créancier et al., 1993). The RGG domain on its own, however, is not a nucleolar localization signal; non-nucleolar proteins fused to an NLS and to the nucleolin RGG domain also fail to localize to nucleoli (Meßmer and Dreyer, 1993; Schmidt-Zachmann and Nigg, 1993).

To test the localization properties of the RGG domain in DmNopp140-RGG, we used the PCR to amplify only the cDNA sequence that encodes amino acids 1–561 of DmNopp140-RGG. The deleted segment extended upstream of the actual RGG domain to include two additional tandem RGG motifs at residues 562-567. The deleted residues (562-688) are italicized in Fig. 1C. We refer to the expressed truncation as DmNopp140ΔRGG. In more than half of the transfected cells, GFP-DmNopp140ΔRGG failed to localize to nucleoli (Fig. 8B). In a few these cells, we could see nucleoli that were barely labeled above background. Instead of localizing to the nucleoli, GFP-DmNopp140ΔRGG localized to the nucleoplasm (Fig. 8B). The nucleoli in these cells appeared morphologically normal (Fig. 8A), and fibrillarin remained localized to these nucleoli (Fig. 8C).

In many of the other transfected cells (approaching half), we observed accumulations of GFP-DmNopp140ΔRGG in large phase-light regions of partitioned
nucleoli (precisely aligned arrows in Fig. 8D and E). Endogenous nucleolin again maintained its localization within the phase-dark regions (Fig. 8F), while endogenous fibrillarin co-localized with DmNopp140ΔRGG in the phase-light regions (not shown). We conclude from these observations that DmNopp140ΔRGG generally fails to associate with morphologically normal phase-dark nucleoli when expressed in low to moderate levels. At higher expression levels, however, GFP-DmNopp140-ΔRGG mimics full-length DmNopp140 in causing nucleoli to partition into phase-light and phase-dark regions. Again fibrillarin and nucleoli partition to phase-light and phase-dark regions, respectively.

Cells expressing GFP-DmNopp140-ΔRGG were counter-stained with anti-human p80 coilin (Fig. 8G-I). In many cells DmNopp140-ΔRGG again distributed throughout the nucleoplasm, but it failed to localize to endogenous CBs that were detected with the anti-coilin (precisely aligned arrowheads in Fig. 8H and I). In cells that contained partitioned nucleoli, however, endogenous coilin again accumulated on the periphery of or within phase-light regions (precisely aligned arrows in Fig. 8G-I). A good example of peripheral localization of endogenous coilin is shown just below the center of Fig. 8I. The top part of this peripheral region appears to contain a spherical CB.

From the observations presented in Figure 8, we conclude that DmNopp140-ΔRGG behaves much like GFP-DmNopp140 in causing nucleoli to partition into phase-light and phase-dark regions. Since DmNopp140-ΔRGG partitions nucleoli in a manner similar to that observed for DmNopp140, the amino terminus and/or the large central domain of DmNopp140 must be responsible for this observed partitioning. The distinctive RGG domain in DmNopp140-RGG may prevent or dampen any propensity of the amino terminus or the central domain within DmNopp140-RGG to partition nucleoli, since partitioning does not occur in cells expressing GFP-DmNopp140-RGG. Finally, DmNopp140-ΔRGG fails to localize to endogenous CBs, but endogenous coilin appears to localize to the phase-light regions containing DmNopp140-ΔRGG. We can conclude that the at least the RGG domain does not prevent DmNopp140-RGG from localizing to CBs.
The reason that the two intact *Drosophila* variants fail to localize to endogenous CBs must lie in sequence or structural differences within the amino terminus or central region as compared to the rat protein that we know localizes to CBs.
DISCUSSION

Two Drosophila Nopp140 Variants

The *Nopp140* gene in *Drosophila melanogaster* maps within polytene segment 79A5. This region is proximal to the centromere (80F) on the left arm of chromosome 3. Transcription is in the direction of the centromere. Genes upstream of *Nopp140* are *eagle* (a steroid hormone receptor/transcription factor) in 79A4, the gene encoding cyclin H (for cell cycle regulation), and conceptual genes CG7407, CG7414, CG7148, all of unknown function. CG7145 maps downstream of *Nopp140*; it encodes a 1-pyrroline-5-carboxylate dehydrogenase-like enzyme. An enhancer-promoter type P-element transposon (EP(3)3138) maps within the promoter of CG7145, approximately 5.5 kbps downstream of the 3' end of *Nopp140* (FlyBase, 1999). In the future, this P-element should allow us to create deficiencies that eliminate the *Nopp140* gene in our efforts to understand Nopp140 function in metazoans.

The *Nopp140* gene encodes a predicted pre-mRNA of 3247 nucleotides that is differentially spliced to produce two transcripts, each encoding a Nopp140 variant. DmNopp140 is the sequence homologue of rat Nopp140. Its overall domain composition and organization is quite similar to the prototypical rat Nopp140, and this is particularly true for its central and carboxy terminal domains. The overall length of the central domain within DmNopp140 (and DmNopp140-RGG), however, is shorter than that in rat Nopp140. Despite its central domain being shorter, DmNopp140 has 14 acidic regions instead of 10 as in rat Nopp140. *Xenopus* xNopp180, on the other hand, has 18 acidic regions (Cairns and McStay, 1995).

The carboxy termini of various Nopp140 proteins serve to best define homology (Meier, 1996; Table I). The entire carboxy terminus of DmNopp140 is 64% identical over a 97 amino acid comparison with the carboxy terminus of the prototypical rat Nopp140.
The individual carboxy sub-domains of DmNopp140 (C\textsubscript{a} and C\textsubscript{b}) are as close in homology to the respective sequences in rat and human Nopp140 as the two Xenopus sub-domains are to the same respective domains in human and rat Nopp140. The fact that individual exons encode these two sub-domains in Drosophila suggests evolutionary constraints on distinct sub-domain function. These functions have yet to be determined precisely, but the properties of these two sub-domains have been described (Isaac et al., 1998). Finally, a consensus cAMP-dependent protein kinase phosphorylation site is present in the conserved carboxy terminus of all Nopp140 homologues (Meier, 1996). The presence of the putative site suggests that Nopp140 is a direct substrate for signal transduction-mediated phosphorylation cascades that may regulate molecular interactions of Nopp140 within nucleoli or Cajal bodies. DmNopp140 contains a similar site (serine 638), but whether this site is utilized \textit{in vivo} is not yet known.

\textit{DmNopp140 Partitions HeLa Cell Nucleoli}

Despite sequence similarities to the mammalian Nopp140 proteins, DmNopp140 causes HeLa cell nucleoli to partition into clearly discernible phase-light and phase-dark regions (Fig. 6A). Partitioning was so severe in some cells (e.g. Fig. 6A and B) that we were often able to identify transfected cells expressing exogenous DmNopp140 simply by using phase contrast microscopy. Endogenous fibrillarin co-localized with DmNopp140 to the phase-light regions (Fig. 6C-E), while endogenous nucleolin localized to the phase-dark regions (Fig. 6F-H). DmNopp140 completely failed to associate with the phase-dark regions. The separation of DmNopp140 from nucleolin is reminiscent of dominant negative effects caused by the transient expression of the carboxy terminal domain of rat Nopp140, the hemagglutinin-tagged NoppC described by Isaac et al. (1998). In their study, exogenous expression of NoppC clearly partitioned fibrillarin and nucleolin in a manner similar, but not identical to our exogenous expression of full length DmNopp140 in HeLa cells. NoppC chased endogenous Nopp140, fibrillarin, and NAP57 out of the nucleoli and into the nucleoplasm. The three proteins did not localize to any nucleolar cap or phase-light regions of the partitioned nucleoli. Nucleolin and UBF, on the other hand,
were not affected by the expression of NoppC; they remained within the phase-dark nucleoli. In our studies, endogenous fibrillarin was “chased” into the phase-light regions of the partitioned nucleoli while nucleolin remained behind in the phase-dark regions.

Chen et al. (1999) described similar nucleolar partitioning in HeLa cells when they over-expressed a carboxy truncation of human Nopp140 (hNopp140N382, containing the first 382 amino acid residues out of the 699 total). In their study, the truncation co-localized with fibrillarin and the large subunit of polymerase I within partitioned nucleolar caps. Chen et al. (1999) proposed that hNopp140N382 caused cap formation by interacting with the large subunit of RNA polymerase I in a dominant negative manner to block rRNA transcription. The authors cited the fact that similar nucleolar caps form when cells are treated with actinomycin D at concentrations known to block rRNA transcription (see the descriptions of actinomycin D-segregated nucleoli by Busch and Smetana, 1970).

It will be interesting to determine if the partitioned nucleoli that we see due to the exogenous expression of DmNopp140 (rather than by actinomycin D) are still functional. Barring any dominant negative effects on transcription caused by DmNopp140, we may be able to determine the active sites for transcription and processing with respect to the phase-light and phase-dark regions. The partitioned nucleoli may allow us to explore possible molecular associations between nucleolar components in vivo. For example, the co-localization of DmNopp140 and endogenous fibrillarin within the phase-light regions is consistent with the observation that Nopp140 associates with fibrillarin in box C/D snRNP complexes (Yang et al. 2000). Fibrillarin associates with U3 and many of the other box C/D snoRNPs. Conversely, nucleolin also interacts with the U3 snoRNP (a box C/D snoRNP) for cleavage within the 5’ external transcribed spacer of pre-rRNA (Ginisty et al., 1998). It remains uncertain if nucleolin interacts with the many other box C/D snoRNPs that selectively methylate pre-rRNA. The clear segregation of endogenous fibrillarin (phase-light region) from endogenous nucleolin (phase-dark region) in HeLa cells that express DmNopp140 suggests that nucleolin and fibrillarin do not associate directly in vivo, despite the fact that they normally co-localize within DFCs. We predict, therefore,
that box C/D snoRNPs (with fibrillarin as antigen) and box H/ACA snoRNPs (with NAP65 as antigen) will co-localize within the partitioned phase-light regions, while nucleolin and B23 remain within the phase-dark regions of partitioned nucleoli in cells expressing DmNopp140. Questions remain: what other nucleolar components reside within the phase-light versus the phase-dark regions? Does U3 localize to both the phase-light and the phase-dark regions?

**DmNopp140-RGG is a Novel Splice Variant**

With its RGG carboxy terminus, DmNopp140-RGG is a novel and striking variant of DmNopp140. Naturally occurring variants of other nucleolar proteins have been described. For instance, *Xenopus laevis* expresses two versions of nucleolin that are encoded by separate genes (Meßmer and Dreyer, 1993). *Xenopus* is pseudo-tetraploid (Kobel and DuPasquier, 1986), and the two modern nucleolin genes may have descended from a common ancestral gene by duplication and divergence. The two nucleolin proteins differ primarily in the number of amino terminal acidic and basic regions. Further, two splice variants of B23 exist 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 the nucleoplasm, but not the nucleolus (Wang et al., 1993). Alternative splicing results in the deletion of the carboxy-terminal 35 residues and the substitution of two additional upstream residues to convert B23.1 to B23.2 (Chang and Olson, 1989, 1990). Interestingly, B23.1, but not B23.2, binds nucleic acids (Wang et al., 1994). As a third example, two isoforms of human Nopp140 have been reported (Pai and Yeh, 1996). The alpha form predominates, but the novel beta form contains a ten amino acid insert in the fourth proline-rich basic region of the central domain. No functional differences are known to exist between the alpha and beta forms of human Nopp140. Compared to these relatively minor differences in the number of alternating acidic and basic regions or the even the deletion variant of the B23, the discovery of DmNopp140-RGG indicates that novel proteins related to the prototypical rat Nopp140 but with strikingly different domains may perform related but non-overlapping functions.
The RGG Domain of DmNopp140-RGG

The extensive glycine and arginine rich (RGG) domain in DmNopp140-RGG is very similar to the RGG domain found in vertebrate nucleolin. Nucleolin's RGG domain is necessary but not sufficient for proper nucleolar localization (Meßmer and Dreyer, 1993; Schmidt-Zachmann and Nigg, 1993; Heine et al., 1993; Créancier et al., 1993). Likewise, DmNopp140-ΔRGG generally fails to associate with nucleoli in roughly half the transfected HeLa cells (Fig. 8B). In these cells, DmNopp140-ΔRGG distributes to the nucleoplasm. Thus, the RGG domain of DmNopp140-RGG appears to be necessary for nucleolar localization at least in mammalian cells. We speculate that the RGG domain in DmNopp140-RGG may share the same interactions with nucleolar proteins (e.g. Cartegni et al., 1996; Bouvet et al., 1998) or RNAs (e.g. Ghisolfi et al., 1992a, b; Hanakahi et al., 2000) that have been attributed to nucleolin by way of its RGG domain.

Besides nucleolin, other nucleolar proteins that contain RGG motifs include (but are not limited to) fibrillarin (Lischwe et al, 1985; Ochs et al., 1985) Gar1, Nopp44/46 in Trypanosoma brucei (Das et al., 1998), the three yeast nucleolar proteins gar2p, Ssb1p, and Nop3p. Ribosomal protein S2 also contains RGG motifs (Suzuki et al., 1991). The arginine residues within RGG motifs of nucleolin and fibrillarin are asymmetrically dimethylated ($N^\epsilon,N^\epsilon$-dimethylarginine), and we strongly suspect that the arginines within the RGG domain of DmNopp140-RGG will be similarly dimethylated. This posttranslational modification does not change the charge of the arginine side chains, but it certainly makes the arginine side chains bulkier and more hydrophobic, in all likelihood to modulate molecular interaction. Arginine methylation within RNA-binding proteins is believed to regulate their protein-RNA (Tao and Frankel, 1992), protein-protein interactions (Liu and Dreyfuss, 1995; Friesen et al., 2001), or perhaps their nucleocytoplasmic shuttling (Shen et al., 1998). The functional significance of RGG methylation in regulating protein-protein interactions is just now coming to light (e.g. Friesen et al., 2001).
The distribution of DmNopp140-ΔRGG to the nucleoplasm or nucleolar phase-light regions in HeLa cells is in sharp contrast to the nucleolar localization observed for rat Nopp140ΔC in monkey COS-1 cells (Isaac et al., 1998). Rat Nopp140ΔC accumulated in the cytoplasm, intact nucleoli, CBs, and in phase-dense nuclear rings referred to as R-rings. Rat Nopp140ΔC did not accumulate in the nucleoplasm (Isaac et al., 1998) as did DmNopp140-ΔRGG. In many other transfected HeLa cells, DmNopp140-ΔRGG localized to the phase-light regions of partitioned nucleoli (e.g. Fig. 8H). Rat Nopp140ΔC did not appear to partition the nucleoli. One possible explanation for these differences between rat Nopp140ΔC and DmNopp140-ΔRGG is that the two deletions may not be directly comparable. The rat Nopp140ΔC was truncated at residue 586, thus deleting the last 118 residues (587-704). This deletion stretched into the back end of the central domain (see Fig. 1A). We deleted residues 561-688 from DmNopp140-RGG, thus leaving behind 10 residues that are unique to DmNopp140-RGG versus DmNopp140 (see Fig. 1C and D). Further work is necessary to determine if this 10 residue segment, now the very carboxy end, is critical for the distribution of DmNopp140ΔRGG to the nucleoplasm as compared to the rat Nopp140ΔC. Alternatively, sequence and structural differences between the amino termini and the central domains of rat Nopp140ΔC and DmNopp140-ΔRGG may account for these differences.

Cajal Bodies Versus Phase-light Regions

Both DmNopp140-RGG and DmNopp140 fail to localize to CBs in the nuclei of Xenopus oocytes and HeLa cells. The best explanation for this discrepancy between the two Drosophila variants and rat Nopp140 may lie in their respective tertiary structures. Direct associations between Nopp140 and coilin have been established using the two-hybrid system and co-immunoprecipitations (Isaac et al., 1998). Specifically, full-length rat
Nopp140 associates well with the amino terminal region (the first 161 amino acids) of coilin. Although deletion of the conserved carboxy terminus of rat Nopp140 (both subregions \( C_a \) and \( C_b \)) greatly diminishes the interaction with coilin, coilin fails to interact with the individual amino terminal domain of Nopp140, the individual central domain, and the individual carboxy domain (Isaac et al., 1998). From these results it appears that the entire Nopp140 sequence is necessary to achieve an association with coilin, perhaps by folding into a particular tertiary structure. Any interaction between Nopp140 and coilin must be reserved for the CBs since p80 coilin does not normally localize to nucleoli (but see below). A particular tertiary structure of Nopp140 that is necessary for coilin interaction must shift when Nopp140 gains access to nucleoli. Such a structural shift in Nopp140 may then permit Nopp140 to interact with other nucleolar components instead of coilin. Phosphorylation of Nopp140 (Meier, 1996) or of coilin (Hebert and Matera, 2000) by CKII may affect their respective structures to thus allow or prevent their intermolecular associations. Perhaps the two *Drosophila* variants are sufficiently different in primary sequence and thus tertiary structure that they fail to interact with endogenous coilin in the CBs of *Xenopus* oocytes and HeLa cells.

Clues as to why the two *Drosophila* variants fail to localize to CBs may be found in the studies describing treacle. Treacle is a human protein related to Nopp140 in that it contains a homologous central repeat domain with 10 repeating acidic and basic regions (Dixon et al., 1997; Wise et al., 1997). Like Nopp140, treacle is highly phosphorylated by CKII, and it localizes to the DFCs of nucleoli. Treacle, however, is distinct from Nopp140 in that it fails to associate with CBs (Isaac et al., 2000). Treacle is the product of the human *TCOF1* gene (Dixon et al., 1997), mutations in which give rise to the Treacher Collins syndrome, an autosomal dominant disorder that affects craniofacial development (Dixon, 1996). Sequence differences between treacle and Nopp140 reside primarily in their amino and carboxy termini, and these difference may account for treacle's failure to localize to CBs (Isaac et al., 2000). Analogous differences between the amino and carboxy termini of rat Nopp140 and the two *Drosophila* variants may therefore account for the failure of the two *Drosophila* variants to localize to CBs.
Despite the fact that the *Drosophila* Nopp140 variants fail to localize to endogenous CBs, endogenous coilin appears to associate with the phase-light regions of partitioned nucleoli in about half the cells expressing DmNopp140 or DmNopp140ΔRGG. In these cells coilin accumulates on its periphery of the phase-light regions, but some phase-light regions also appear to contain small amounts of p80 coilin within their interiors. Previous observations established that coilin can localize to intranucleolar structures under certain conditions. For example, coilin localizes to nucleoli within liver cells and brown adipocytes of hibernating hazel dormice (Malatesta et al., 1994) and within certain human breast cancer cells (Ochs et al., 1994). Simple over-expression of GFP-tagged full-length coilin causes its accumulation in spherical structures within nucleoli, while two GFP-coilin truncations, GFP-coilin(1-248) GFP-coilin(1-315), co-localize with fibrillarin and Nopp140 within the DFCs of HeLa nucleoli (Hebert and Matera, 2000). Several other results suggest that phosphorylation plays an important role in coilin's localization to nucleoli. First, endogenous coilin and U2 snRNPs can redistribute to intranucleolar structures within HeLa cells when Ser/Thr dephosphorylation is inhibited by 10 nM okadaic acid (Lyon et al., 1997, Sleeman et al., 1998). Second, a dominant mutant of coilin (S202D) that mimics constitutive phosphorylation localizes to intranucleolar structures, and causes the disassembly of all other extranucleolar CBs (Lyon et al., 1997, Sleeman et al., 1998). Conversely, the S184A mutation in human coilin mimics the dephosphorylated state, and the mutant co-localized with fibrillarin in 30% of the cells expressing the mutation (Hebert and Matera, 2000). Finally, a cryptic nucleolar localization signal within coilin may be differentially exposed depending on the state of coilin phosphorylation (Hebert and Matera, 2000). In many of these cases coilin associated with a defined spherical structure within the nucleoli, structures that are reminiscent of the FBs. We refrain from identifying the phase-light regions as swollen FCs since the phase-light regions contain exogenous DmNopp140 and endogenous fibrillarin, both of which are considered to be DFC markers. It is interesting to note that fibrillarin is a component of both CBs and the phase-light regions of the partitioned nucleoli. Nucleolin, on the other hand, is not a component of either the somatic cell CBs or the phase-light regions. Future work with antibodies directed
against known FC and DFC markers should better define the molecular composition of the phase-light regions and how these phase-light regions relate to CBs.

Finally, we also predict that molecular interactions (and functions) of DmNopp140-RGG versus DmNopp140 will prove different within nucleoli. The molecular cytology and molecular genetics available in Drosophila should give us ample opportunities to explore these differences. For example, we fully expect DmNopp140 to interact with the minifly gene product, Nop60B (Phillips et al., 1998) in Drosophila. But will the same interaction occur with DmNopp140-RGG? Drosophila also allows us to explore developmental and tissue-specific expression patterns of the two variants. What we learn about the molecular interactions of DmNopp140 versus DmNopp140-RGG and the expression patterns of these two splice variants, should advance our knowledge of novel and traditional nucleolar functions.

ACKNOWLEDGEMENTS

We thank Drs. John Tower and Allan Spradling (Department of Embryology, Carnegie Institution of Washington, Baltimore, MD) for an aliquot of their stage 10 egg chamber cDNA library. We also thank Dianne Stewart and Allan Spradling for the Schneider II cells. We thank Drs. Zheng’an Wu and Joe Gall (Department of Embryology, Carnegie Institution of Washington, Baltimore, MD) for the R288 rabbit serum directed against human coilin and for insightful discussions regarding CBs. We thank Dr. Benigno (Ben) Valdez (Baylor College of Medicine, Houston, TX) for antibodies directed against human nucleolin and Dr. Robert Ochs (Scripts Research Institute, La Jolla, CA) for the anti-human S4 anti-fibrillarin antibody. This work was supported by NSF grant MCB-9727917.

REFERENCES


FIGURE LEGENDS

Figure 1. Sequence comparisons of DmNopp140-RGG and DmNopp140. (A) Acidic (black boxes) and basic (white boxes) regions alternate within the central domains of both Drosophila variants. This alternating pattern is similar to that within rat Nopp140. DmNopp140 is the Drosophila homologue of mammalian Nopp140, while DmNopp140-RGG is unique due to its carboxy terminal RGG domain. (B) Expressed sequence tags exist for both Drosophila variants. Restriction maps allowed us to predict digestion patterns, and thus to verify the clones. The purified cDNA insert that encodes conceptual transcript 22833 (DmNopp140) was left undigested (lane 1), incubated with PvuI that failed cut as predicted (lane 2), or digested with AvaI 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 PvuI that generated two predicted fragments (lane 5), or incubated with AvaI that failed to cut as expected (lane 6). The digestion patterns confirmed the existence of two separate cDNA clones. (C) The deduced amino acid sequence of DmNopp140-RGG. CT22845 consists of three exons. The first encodes amino acids 1-34, the second encodes 35-551, and the third encodes 552-688. 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 upstream RGG motifs reside at 562-567. (D) The deduced amino acid sequence of DmNopp140. CT22833 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 is 64% identical to the carboxy terminus of rat Nopp140. A highly conserved serine residue within this terminus (residue 610 in the black box) is a putative substrate site for cAMP dependent protein kinase.

Figure 2. In vitro MPF and CKII phosphorylation. 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 right panel is the corresponding autoradiogram. Lanes A and A’: protein incubated with MPF and $[^{32}\text{P}]\text{-gamma-labeled ATP}$. Lanes B and B’: protein incubated with CKII and gamma-labeled ATP. Lanes C and C’: purified protein incubated with gamma-labeled ATP, but 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 (arrow).

**Figure 3.** *Drosophila* Schneider II cells transfected to express GFP-DmNopp140-RGG and GFP-DmNopp140. GFP-DmNopp140-RGG localized to nucleoli (arrow in panels A and B). No other nuclear bodies were apparent (panel B). GFP-DmNopp140 also localized to nucleoli (panels C and D). No other nuclear bodies were apparent (panel D).

**Figure 4.** DmNopp140-RGG and DmNopp140 localize to the DFCs of *Xenopus* oocyte nucleoli, but they fail to localize to the sphere organelles. *Xenopus* oocytes were injected with synthetic transcripts that encoded either GFP-DmNopp140-RGG (panels A-E) or GFP-DmNopp140 (panels F-J). Panels A, D, F, and I are phase contrast images showing a few of the approximately 1000 nucleoli in each oocyte nucleus. GFP-DmNopp140-RGG (panel B) and GFP-DmNopp140 (panel G) both localized within the DFC regions of the nucleoli, specifically within a sub-region immediately surrounding the FCs. These FCs contained detectable amounts of the nucleolar DNA as demonstrated by DAPI staining (panels C and H). Some nucleoli typically contained multiple FCs and DFCs (e.g. panels G and H). Both GFP-DmNopp140-RGG (panels D and E) and DmNopp140 (panels I and J) failed to localize to the large sphere organelles (arrows) that are homologous to Cajal bodies. Calibration bar is 20 $\mu$m for all panels.

**Figure 5.** DmNopp140-RGG localizes to intact, phase-dark nucleoli but fails to localize to endogenous Cajal bodies in transfected HeLa cells. **Panels A and B:** GFP-DmNopp140-
RGG (panel B) localized to nucleoli in a punctate staining pattern. **Panels C-E:** Exogenous GFP-DmNopp140-RGG (panel D) co-localized with endogenous fibrillarin (panel E) that was detected with the S4 anti-human fibrillarin antibody and a rhodamine-conjugated secondary antibody. Arrowheads in panels C-E point to nucleoli that contain relatively large amounts of DmNopp140-RGG while arrows in panels C-E point to nucleoli that maintain a punctate staining pattern with moderate amounts of DmNopp140-RGG. **Panels F-H:** Exogenous GFP-DmNopp140-RGG (panel G) co-localized with endogenous nucleolin (panel H) that was detected with an anti-nucleolin antibody and a rhodamine-conjugated secondary antibody. Arrows in panels F-H point to nuclear bodies that fail to stain with DmNopp140-RGG or anti-nucleolin. Conversely, all labeled structures evident in panels B, D, and G corresponded well to phase-dark nucleoli in panels A, C, and F, respectively. **Panels I-K:** Exogenous GFP-DmNopp140-RGG (panel J) again localized to intact phase-dark nucleoli, but it failed to co-localize with coilin in endogenous CBs as detected by the anti-human coilin antibody, R288 (panel K). Precisely aligned arrows point to endogenous CBs. Calibration bar is 20 µm for all panels.

**Figure 6.** DmNopp140 partitions HeLa nucleoli into phase-light and phase-dark regions, and it also fails to localize to endogenous CBs. **Panels A and B:** GFP-DmNopp140 localized exclusively to the phase-light regions of the nucleoli (panel B). The phase-dark regions (arrows in panels A and B) were completely devoid on GFP-DmNopp140. **Panels C-E:** GFP-DmNopp140 (panel D) co-localized with endogenous fibrillarin (panel E) in the phase-light regions of the partitioned nucleoli. The S4 anti-human fibrillarin antibody and a rhodamine-conjugated secondary antibody detected the endogenous fibrillarin (panel E). Precisely aligned arrows point to phase-dark regions. **Panels F-H:** GFP-DmNopp140 (panel G) again localized to the phase-light regions, while endogenous nucleolin (panel H) localized to the phase-dark regions (arrows in panels F-H). An anti-human nucleolin antibody and a rhodamine-conjugated secondary antibody detected the endogenous nucleolin (panel H). **Panels I-K:** GFP-DmNopp140 localized to the phase-light regions of partitioned nucleoli, and precisely matched arrows show the phase-dark regions. The anti-human coilin antiserum, R288, detected endogenous CBs (arrowheads in panels K). GFP-
DmNopp140 failed to localize to endogenous CBs, but endogenous coilin appeared to accumulate on the periphery of or within the phase-light regions (panel K). Calibration bar is 20 µm for all panels.

**Figure 7.** 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 on fluorescence signals. Both proteins co-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). Nucleoli appeared segregated by phase contrast microscopy (panel D), yet both proteins co-localized to the phase-light regions. **Panels G-I:** GFP-DmNopp140 (panel H) was under-expressed relative to RFP-DmNopp140-RGG (panel I). Nucleoli appeared morphologically normal by phase contrast microscopy (panel G), yet both proteins co-localized. Calibration bar is 20 µm for all panels.

**Figure 8.** Expression of the carboxy terminal truncation, DmNopp140ΔRGG, in transfected HeLa cells. **Panels A-C:** GFP-DmNopp140ΔRGG failed to localize to intact phase-dark nucleoli. The cells were counter-stained with the S4 anti-human fibrillarin antibody and a rhodamine-conjugated secondary antibody (panel C). Most of the endogenous fibrillarin remained associated with the nucleoli. **Panels D-F:** In approximately half the transfected cells GFP-DmNopp140ΔRGG caused nucleoli to partition into phase-light (precisely matched arrows in panels D-F) and phase-dark regions, reminiscent of the effects caused by GFP-DmNopp140. The cells were counter-stained with anti-human nucleolin and a rhodamine conjugated secondary antibody (panel F). Nucleolin localized only within the phase-dark regions that were completely devoid of GFP-DmNopp140ΔRGG. **Panels G-I:** Transfected HeLa cells again expressed GFP-
DmNopp140ΔRGG (panel H). Some cells displayed a diffuse distribution of GFP-DmNopp140ΔRGG while other cells displayed partitioned nucleoli with DmNopp140ΔRGG localizing to the phase-light regions (precisely aligned arrows in panels G-H). The cells were counter stained with the R288 anti-serum directed against human coilin (panel I). Arrowheads show endogenous CBs that did not contain GFP-DmNopp140ΔRGG. Conversely, arrows show phase-light regions that appear to contain endogenous coilin. Calibration bar is 20 μm for all panels.
Figure 1A

DmNopp140-RGG

RGG

DmNopp140

Rat Nopp140
Figure 1B
### Figure 1C

**DmNopp140-RGG**

<table>
<thead>
<tr>
<th>1</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTDLLKIADA</td>
<td>IVLEYLQSKD</td>
<td>KNLAKVFQK</td>
<td>TKAASVAKSS</td>
<td>PKLSEILQFY</td>
<td>QTKSPKKIPA</td>
<td>IKATACDSSE</td>
<td>DSDSDSES</td>
</tr>
<tr>
<td>APKKPATAPA</td>
<td>LTNGKAVKKA</td>
<td>ASSSTSEDSDS</td>
<td>EEKKPAAKA</td>
<td>TPAKAVGKKA</td>
<td>KSSSSEDSSSE</td>
<td>EEAPKKAAPV</td>
<td>KAPPKKAAPA</td>
</tr>
<tr>
<td>KKVESSSDDS</td>
<td>SSEEPEAKPA</td>
<td>VKATTTKVAP</td>
<td>AKKADSSSEE</td>
<td>SSSDEETKPA</td>
<td>AKPAAKAAPA</td>
<td>KKAASSSEES</td>
<td>DSDDEPAAKK</td>
</tr>
<tr>
<td>PAVQPAKPA</td>
<td>PKAAASSSED</td>
<td>SSESEEVKPA</td>
<td>AKSAALKAPA</td>
<td>KKAASSSDDS</td>
<td>SSEDEAPKKA</td>
<td>ATLAIFISK</td>
<td>APTKKAADSSI</td>
</tr>
<tr>
<td>EDSSSEDDAP</td>
<td>KKVPAPAKA</td>
<td>TP AKAIAPKKA</td>
<td>SSSDDSSSEE</td>
<td>APKKAAPAKA</td>
<td>TPAKATPAKAK</td>
<td>AASSDDSSSE</td>
<td>EEAAPKKAAP</td>
</tr>
<tr>
<td>AKATAKKAK</td>
<td>SSESDDSDE</td>
<td>EEAAPKPAKA</td>
<td>AVAKAASSSED</td>
<td>SSSSEDEKPA</td>
<td>KAAPKALAKS</td>
<td>KKAASSSDS</td>
<td>SSDEETPAVK</td>
</tr>
<tr>
<td>PAVKTTAAPA</td>
<td>KKADDSSES</td>
<td>DSGEEEVKPA</td>
<td>PNSATNGNEK</td>
<td>TAQKRKFSGG</td>
<td>DQDEATPNKK</td>
<td>YNNFVKSGEQ</td>
<td>EENDFKKHN</td>
</tr>
<tr>
<td>GGGRGGGGGF</td>
<td>SGRPDRSTWE</td>
<td>TKNFNGEGGG</td>
<td>DGGGFKKIGD</td>
<td>RKSFGDFDNN</td>
<td>GGGGGGGGGRG</td>
<td>GGGGGGGGGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Figure 1D

**DmNopp140**

<table>
<thead>
<tr>
<th>1</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTDLLKIADA</td>
<td>IVLEYLQSKD</td>
<td>KNLAKVFQK</td>
<td>TKAASVAKSS</td>
<td>PKLSEILQFY</td>
<td>QTKSPKKIPA</td>
<td>IKATACDSSE</td>
<td>DSDSDSES</td>
</tr>
<tr>
<td>APKKPATAPA</td>
<td>LTNGKAVKKA</td>
<td>ASSSTSEDSDS</td>
<td>EEKKPAAKA</td>
<td>TPAKAVGKKA</td>
<td>KSSSSEDSSSE</td>
<td>EEAPKKAAPV</td>
<td>KAPPKKAAPA</td>
</tr>
<tr>
<td>KKVESSSDDS</td>
<td>SSEEPEAKPA</td>
<td>VKATTTKVAP</td>
<td>AKKADSSSEE</td>
<td>SSSDEETKPA</td>
<td>AKPAAKAAPA</td>
<td>KKAASSSEES</td>
<td>DSDDEPAAKK</td>
</tr>
<tr>
<td>PAVQPAKPA</td>
<td>PKAAASSSED</td>
<td>SSESEEVKPA</td>
<td>AKSAALKAPA</td>
<td>KKAASSSDDS</td>
<td>SSEDEAPKKA</td>
<td>ATLAIFISK</td>
<td>APTKKAADSSI</td>
</tr>
<tr>
<td>EDSSSEDDAP</td>
<td>KKVPAPAKA</td>
<td>TP AKAIAPKKA</td>
<td>SSSDDSSSEE</td>
<td>APKKAAPAKA</td>
<td>TPAKATPAKAK</td>
<td>AASSDDSSSE</td>
<td>EEAAPKKAAP</td>
</tr>
<tr>
<td>AKATAKKAK</td>
<td>SSESDDSDE</td>
<td>EEAAPKPAKA</td>
<td>AVAKAASSSED</td>
<td>SSSSEDEKPA</td>
<td>KAAPKALAKS</td>
<td>KKAASSSDS</td>
<td>SSDEETPAVK</td>
</tr>
<tr>
<td>PAVKTTAAPA</td>
<td>KKADDSSES</td>
<td>DSGEEEVKPA</td>
<td>PNSATNGNEK</td>
<td>TAQKRKFSGG</td>
<td>DQDEATPNKK</td>
<td>YNNFVKSGEQ</td>
<td>EENDFKKHN</td>
</tr>
<tr>
<td>GGQIDVGVNS</td>
<td>IKFD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

56
Figure 2
Figure 3
Figure 4
Figure 5
Figure 7
Figure 8