minifly, A Drosophila Gene Required for Ribosome Biogenesis
Ennio Giordano, Ivana Peluso, Stefania Senger, and Maria Furia
Dipartimento di Genetica, Biologia Generale e Molecolare, Università di Napoli Federico II, I-80134 Napoli, Italy

Abstract. We report here the genetic, molecular, and functional characterization of the Drosophila melanogaster minifly (mfl) gene. Genetic analysis shows that mfl is essential for Drosophila viability and fertility. While P-element induced total loss-of-function mutations cause lethality, mfl partial loss-of-function mutations cause pleiotropic defects, such as extreme reduction of body size, developmental delay, hatched abdominal cuticle, and reduced female fertility. Morphological abnormalities characteristic of apoptosis are found in the ovaries, and a proportion of eggs laid by mfl mutant females degenerates during embryogenesis. We show that mfl encodes an ubiquitous nucleolar protein that plays a central role in ribosomal RNA processing and pseudouridylation, whose known eukaryotic homologues are yeast Cfb5p, rat NA P57 and human dyskerin, encoded by the gene responsible for the X-linked dyskeratosis congenita disease. mfl genetic analysis represents the first in vivo functional characterization of a member of this highly conserved gene family from higher eukaryotes. In addition, we report that mfl hosts an intron encoded box H/ACA snoRNA gene, the first member of this class of snoRNA's identified so far from Drosophila.

Key words: Drosophila • rRNA • ribosome • nucleolus • snoRNA

In eukaryotic cells, synthesis, maturation and modification of rRNA take place in the nucleolus, and RNP composed of a variety of nucleolar proteins and small nucleolar RNA's (snoRNA's) are known to be responsible for these essential cellular processes (reviewed by Melese and Xue, 1995). Almost 100 different snoRNA's species have been identified so far in yeast and mammalian cells. Recently, it became evident that most of these snoRNA's can be classified into two major distinct families, each defined by common associated proteins and by the presence of conserved sequences, designated as either C/D or H/ACA boxes (reviewed by Balakin et al., 1996; Smith and Steitz, 1997). The C and D box-containing snoRNA's display extensive sequence complementarity to conserved rRNA regions and are associated with a conserved nucleolar protein, fibrillarin or, in yeast, with the fibrillarin homologue Nop1p (reviewed by Bachellerie and Cavaille, 1997). Some fibrillarin-associated snoRNA's are required for rRNA processing, but most of them function as a guide in site-specific ribose methylation of rRNA (Kiss-Laszlo et al., 1996; Nicoloso et al., 1996).

Members of the other large class of snoRNA's share H and ACA elements and have only short rRNA complementary motifs, brought together by a conserved stem-loop secondary structure (Ganot et al., 1997b). This structure, composed of two hairpins connected and followed by short single-stranded regions containing the H and ACA elements, directs the site-specific pseudouridylation event with the short (5–9 nucleotide [nt]) regions of snoRNA/rRNA complementarity flanking both sides of the target site (Ganot et al., 1997a; Ni et al., 1997).

In yeast, members of the box H/ACA class of snoRNA's are specifically associated with two essential nucleolar proteins, Gar1p and Cbf5p (Balakin et al., 1996; Ganot et al., 1997b; Lafontaine et al., 1998). Gar1p, a glycine-arginine-rich protein required for accumulation of mature 18S rRNA (Balakin et al., 1996; Girard et al., 1992) and for rRNA pseudouridylation (Bousquet-Antonelli et al., 1997), is thought to play a crucial role in structuring box H/ACA sno-RNPs and favoring association of H/ACA snoRNA's to the pre-rRNA. In a two hybrid yeast assay, Gar1p interacts with Cbf5p which, in turn, coprecipitates with box H/ACA snoRNA's and is required for their stability (Lafontaine et al., 1998). Interestingly, Cbf5p is the yeast member of a highly conserved protein family that includes homologues from at least 18 organisms. A mong eukary-

Dear Ennio, Ivana, Stefania, and Maria,

I’m writing from the Rockefeller University Press to inquire about the publication of your paper titled “minifly, A Drosophila Gene Required for Ribosome Biogenesis” in The Journal of Cell Biology. Your work explores the genetic, molecular, and functional characterization of the Drosophila melanogaster minifly (mfl) gene. Genetic analysis shows that mfl is essential for Drosophila viability and fertility. While P-element induced total loss-of-function mutations cause lethality, mfl partial loss-of-function mutations cause pleiotropic defects, such as extreme reduction of body size, developmental delay, hatched abdominal cuticle, and reduced female fertility.

Morphological abnormalities characteristic of apoptosis are found in the ovaries, and a proportion of eggs laid by mfl mutant females degenerates during embryogenesis. You show that mfl encodes an ubiquitous nucleolar protein that plays a central role in ribosomal RNA processing and pseudouridylation, whose known eukaryotic homologues are yeast Cfb5p, rat NA P57 and human dyskerin, encoded by the gene responsible for the X-linked dyskeratosis congenita disease. mfl genetic analysis represents the first in vivo functional characterization of a member of this highly conserved gene family from higher eukaryotes.

In addition, you report that mfl hosts an intron encoded box H/ACA snoRNA gene, the first member of this class of snoRNA's identified so far from Drosophila.

I am interested in learning more about the significance of your findings and the implications for further research in the field of ribosome biogenesis. Could you provide some insights or guidance on how to proceed with the follow-up research?

Thank you for your time and consideration.

Sincerely,

[Your Name]
otes, genetic analysis has so far been restricted to two members of this family: the yeast Cbf5 (Jiang et al., 1993) and the Drosophila minifly gene, whose mutations cause the X-linked dyskeratosis congenita disease (Hess et al., 1998). Whereas little information is available on human dyskerin, Cbf5 and its rat homologue, NA P57, are known to be proteins with prevalent nucleolar localization (Cadwell et al., 1997; Mieir and Blobel, 1994). However, whereas it has recently been proposed that NA P57 may be involved in nucleocytoplasmic shuttling (Mieir and Blobel, 1994), the yeast protein has been shown to be required for transcription, processing and efficient rRNA pseudouridylation (Cadwell et al., 1997; Lafontaine et al., 1998). This last finding raises the possibility that Cbf5 might act as eukaryotic rRNA pseudouridinylase, a role originally suggested (Cadwell et al., 1997) by its homology with E. coli Trub and P35 synthase. Considering the multiple, essential functions played by Cbf5 in yeast cells, the definition of the roles played by members of this family in multicellular organisms appears to be a relevant issue that deserves extensive investigation.

Here we describe the cloning of the Drosophila member of the Cbf5/Nap57/DK C1 gene family, that we called minifly (mfl), and report a detailed genetic, molecular, and functional analysis of its expression. With the isolation and the characterization of mfl mutants reported in this paper, we provide the first animal model system for the study of the molecular basis of the DK C1 human disease. Our data also reveal that mfl has an intriguing molecular organization, hosting an intron-encoded box H/ACA snoRNA that represents the first member of this class thus far described in Drosophila. We named this RNA snoH1 and suggest that it may be functionally equivalent to the human U7 snoRNA.

Materials and Methods

P-Element Mutagenesis/Enhancer-Trap Schemes, P-Cytogenetic Mapping, Construction of Transformed Lines, and Lethal Phase Analysis

The genetic markers and chromosomes used for mutagenesis and mapping are described in Lindsay and Zimm (1992). Most stocks were from the Bloomington Drosophila Stock Center, while the l(2)k06308 and l(2)01361 strains were provided by the Berkeley Drosophila Genome Project Stock Center. The mfl allele was isolated in a small-scale P-element mutagenesis screen performed essentially according to the “reversion jumping” scheme (Towler et al., 1993). In our experiments, tocl201361, a lethal P(γlacz, ry)1 (O’Kane and Gehring, 1987) insertion at the tooth locus, was mobilized by the P(γlacz, ry)1(l2-D-3)998 element (Laski et al., 1986; Robertson et al., 1988) as a source of transposable. Males carrying both the l(2)01361 marker were collected from mfl mutant background. Lethal phase analysis was performed according to Porter (1987).

Cloning Techniques

Basic cloning techniques, DNA and RNA extraction, manipulation and labeling, screening and sequencing techniques were carried out according to Sambrook et al. (1989).

RNA and Protein Analysis

For Northern blot analysis, 5 μg of poly(A+) or 10 μg of total RNA were electrophoresed and transferred to Hybond-NX filters for hybridization. The 5′ end of SnoH1 RNA was determined by primer extension analyses, using 50 μg of total RNA together with primers complementary to nucleotides 96-135 and 149-189 of the fourth mfl intron. rRNA processing was studied by [3H]uridine (1 mCi/ml, 22.4 Ci/nmol) incorporation in Drosophila larval tissue. Ater 48 h, total RNA was extracted and analyzed by agarose electrophoresis followed by fluorography, as described by Tollervey (1987).

In RNA northern blot analyses, probe I corresponds to oligonucleotide 5′-GTTAAACATCTTTTATAGGTTGCCCAAGGCCCAACC3′; probe II to oligonucleotide 5′-CAACATTTTTGCTGCTATACTCAATTTCTCCTAAATAAGT3′; probe III to oligonucleotide 5′-CTATTTCCGAAATCTTTAATAAGGAACATTTCTAGATG3′. Mapping of Drosophila ribosomal pseudouridines was performed essentially as described by Bakin and Efremov (1993) using as primer the oligonucleotides: 5′-AATCTAAGTTGGCCTAAGTTTTGCGAACAACGGATAAC3′ for 185 U1820, 1821, and 1822; 5′-GCGTCGTTATACATGTGGCCCTTAACTGCTTC-3′ for 185 U 830/831, 840, 841, and 885; 5′-CATTCTACGTGGGAATCTTTCCGCAA3′ for 285 U 2422, U 2444, and U 2499. Western blots were analyzed with a 1:1,000 dilution of an affinity-purified rabbit anti-MFL antibody, kindly provided by S. Poole (University of California, Santa Barbara, CA).

In Situ Analysis

Whole mount ovaries in situ hybridization, using single-stranded DIG-labeled probes, obtained by PCR, and immunohistochemical staining of ovaries were performed essentially as described in Ashburner (1989). The rabbit primary anti-MFL antibody, kindly provided by S. Poole, was diluted 1:400 and detected with a biotin-conjugated secondary antibody and a horseradish peroxidase-biotin–avidin complex (ABC Elite kit; Vector Labs).

Computer Analysis

Sequence comparisons were performed using the BLAST search algorithms available at the National Center for Biotechnology Information Web pages: multiple alignments were performed using the CLUSTAL and BLASTSHADE programs. The snnH1 RNA putative secondary structure was established using the MFOLD program.

Results

Isolation, Genetic, and Phenotypical Analysis of mfl Mutants

The first minifly allele, mfl1, was isolated in our laboratory in the course of a P-element mutagenesis screen on the second chromosome (see Materials and Methods) as a via-

The Journal of Cell Biology, Volume 144, 1999

1124
ble, recessive mutation causing a variety of phenotypic abnormalities. The mfl1 pleiotropic phenotype included an extreme reduction of body size (Fig. 1, a and b), developmental delay, essentially due to a 4–5-d prolongation of the larval life, defects in the abdominal cuticle (Fig. 1 c), strong reduction in the length and thickness of abdominal bristles, and reduced female fertility. Most traits of the mfl1 phenotype largely overlapped those caused by the Drosophila Minute (Kay and Jacobs-Lorena, 1985), mini (Procunier and Tartof, 1975) or bobbed mutations (Boncinelli et al., 1972) that affect, respectively, the synthesis of ribosomal proteins, 5S, or 18S and 28S rRNAs. This similarity suggested for mfl a possible role in ribosome biogenesis, encouraging us to attempt the molecular cloning of the gene.

mfl1 mutation was caused by a single P-element insertion, which, by in situ hybridization of a P-specific probe to salivary gland polytene chromosomes of mfl1 heterozygous larvae (Fig. 1 d), was mapped on the chromosome arm 2R, at the 60B-60C polytene subdivisions boundary. Given that wild-type revertants could be recovered from dysgenic crosses after precise excision of the element (see Materials and Methods), mfl1 mutation appeared to be directly caused by this single PZ insertion. Complementation analysis assigned the gene to the region covered by the Df(2R)Px4 deficiency. Among a number of P-induced lethal mutations recovered by Török et al. (1993) and subsequently deposited as part of the Berkeley Drosophila Genome Project, five mapped at the 60B-60C polytene subdivisions boundary. These mutations were all tested in a complementation analysis, by crossing each of them to mfl1 heterozygous flies. Two lines, l(2)k05318 and l(2)k06308, yielded transheterozygous flies with a strong mfl phenotype at the expected ratio, leading us to conclude that they belonged to the mfl complementation group and represented lethal mfl alleles. Accordingly, these lines were renamed, respectively, mfl05 and mfl06.

Previous cytological mapping by the Berkeley Drosophila Genome Project assigned these two mfl alleles to the poly-
tene interval 60B11-C2, in good agreement with our results. By lethal phase analysis (Fletcher et al., 1995; see below) we observed that mfl5 homozygous die mainly as first instar larvae, while most of the mfl6 animals die later, either as second or mainly as young third-instar larvae. Both mfl5 and mfl6 animals fail to increase their size as compared with their wild-type heterozygous siblings and survive for an additional 4–5 d as first or third instar larvae, respectively.

Since a feature of the mfl pleiotropic phenotype was represented by reduced female fertility, we looked at the structure of mutant ovaries. Morphological abnormalities were often observed, with some of the egg chambers beginning to degenerate beyond approximately stage 7 (according to King, 1970) of oogenesis (Fig. 2 c, see brackets). In the degenerating egg chambers, fragmented or condensed nurse cell nuclei with irregular shape are frequently found (Fig. 2 c, boxed). These observations raise the possibility that apoptotic cell death may occur in mfl abnormal ovaries. This possibility was investigated by staining egg chambers with acridine orange (AO). AO is a vital dye that is known to selectively stain apoptotic cells in insects (Spreij, 1971; A brasms et al., 1993) and has successfully been used to study the distribution of apoptosis in Drosophila ovaries (Foley and Cooley, 1998). In our experiments, wild-type ovaries exhibit a diffuse green fluorescence (Fig. 2 b), whereas highly fluorescent yellow spots are detected in mfl3 degenerating egg chambers (Fig. 2 d). These yellow spots are known to correspond to apoptotic, AO highly positive nuclei (Foley and Cooley, 1998), thus confirming the occurrence of apoptosis in mfl ovaries.

A consequence of the gonadic abnormalities observed, mfl homozygous females lay a reduced number of mature eggs, and ~15% of the embryos produced failed to hatch. Such degenerating embryos show asynchronous and atypical development, invariably accompanied by diffuse apoptotic cell death (data not shown). Many mutations causing partial loss-of-function of vital genes interfere with the proper development of the egg, causing female sterility. An important aspect of the results described was the observation that the mfl mRNA subforms are expressed in a temporally and spatially regulated manner within the ovary. These observations suggest that the mfl gene may play a role in the regulation of apoptosis during oogenesis.

Molecular Organization, Coding Properties, and Developmental Expression Profile of the Minify Gene

The genomic region adjacent to the PZ transposon was cloned from the mfl3 stock by plasmid rescue (Wilson et al., 1989) and used to isolate the sequences encompassing the PZ insertion site. Genomic probes spanning a region of ~4 kb surrounding PZ insertion identified on Northern blots of poly(A)^+ RNA two main transcripts of 1.8 and 2.0 kb in length, whose expression was affected in each mfl mutant line (see next section). While the 1.8-kb species was constitutively expressed throughout the life cycle, the 2.0-kb RNA was specifically found in adult female and embryonic RNA preparations, in which a further transcript of ~4.0 kb was also occasionally detected (Fig. 3 b). However, no cDNA representative of this mRNA subform was isolated after extensive screening of an adult female cDNA library, so that it remains unclear whether it actually derives from the mfl gene. In contrast, several cDNAs representative of the 1.8 and 2.0 kb were isolated from adult female and larval libraries. The longest cDNA of each class, respectively, of 1,833 and 2,034 bp, including the poly(A) tail, represented almost full-length transcripts and allowed us to define the mfl gene structure by Southern blot hybridization and alignment with nucleotide sequence of the genomic region. In each mfl mutant line, a copy of P was inserted at the 5' common end of 1.8- and 2.0-kb transcription units: in mfl6, the insertion site was mapped 18 nt upstream from the 5' end of the longest cDNA obtained, in mfl118 nt downstream, within the 5' leader sequence, while in mfl9 the insertion occurred within the first intron of the gene (see Fig. 3 a). The 1.8- and 2.0-kb mRNA subforms share a common coding region and differ from each other only at their alternatively spliced 3' untranslated region, where two additional exons (7 and 8) are specifically included in the 2.0-kb mRNA. When used on northern blots, a probe derived from these two exons (probe 2, depicted below the genomic map) detects exclusively the 2.0-kb subform, specifically present in embryos and adult female RNA's (Fig. 3 c). Hybridization of this probe to whole mount preparations of wild-type ovaries reveals that the female transcript accumulates in germ line cells from the early germarial till last oogenesis stages (Fig. 3 d). We then followed the accumulation profile of both mfl mRNA's during embryogenesis by developmental northern blot analysis of carefully synchronized embryos. As depicted in Fig. 3 e, both mfl mRNA's are detected in very early, 0–2 h embryos. However, while the zygotic 1.8-kb mRNA persists at later stages, the level of female transcript drops subsequently, and becomes very low in 4–6 h embryos. This developmental pattern is very similar to that of other stable maternally supplied RNA's, which persist from early stages up to gastrulation.

The mfl open reading frame (ORF), identically present in both mRNA subforms, encodes a predicted protein of 508 amino acids with a calculated molecular mass of 56 kDa. Database searches revealed that this protein belongs to the Cbf5p/NA P57/dyskerin family (Fig. 3 f). The MFL polypeptide shows a significant degree of conservation to other members of the family, particularly with the two very similar rat and human proteins (66% identity, 79% similarity to human dyskerin). The conservation increases remarkably within several specific domains, strongly underlining that their function has been preserved during evolution. As depicted in Fig. 3 f, total identity exists among Drosophila and human proteins within the two TruB motifs which have homology with bacterial and yeast tRNA pseudouridine synthases (Hess et al., 1998). A repeated hydrophobic domain, possibly involved in the nucleo-cytoplasmic shuttling postulated for the rat protein is also highly conserved. This domain is immediately followed by a block of >20 amino acids having a central tyr that is identical in Drosophila, rat, and human proteins. A though no function has been suggested so far for this domain, its conservation suggests that it might play a rele
vant role in protein activity. Within the tyr domain, we noticed a RX-x(2,3)-DE-x(2,3)-Y central core motif highly conserved among uracil-DNA glycosylases from different organisms as part of the rigid uracil-binding pocket (Up) present in these repair enzymes. The tyrosine residue has been shown to be directly involved in uracil recognition (Kavli et al., 1996; Slupphaug et al., 1996). By analogy, it is reasonable to suggest that the highly conserved tyr motif might play a similar role in uracil recognition. A highly charged lysin-rich COOH-terminal region containing a nuclear localization signal is found in MFL, as in NAP57 and dyskerin, and the NH2-terminal nuclear localization signal observed in rat and human proteins is also preserved. Finally, it is interesting to note that all five missense mutations thus far identified in DKC patients fall into regions that are conserved between the human and the Drosophila gene (see positions of asterisks in Fig. 3f).

Figure 3. Molecular characterization of the minifly gene. (a) Restriction map of the genomic region encompassing the minifly gene (S, SalI; E, EcoRI; B, BamHI). Genomic DNA sequence can be obtained from GenBank (accession number AF097634). On the top, position of P-element insertions. Below, organization of the 1.8- and 2.0-kb mfl transcription units. Exonic regions spanned by mfl ORF are depicted in black. Nucleotide sequence of the mfl maternal transcript can be obtained from GenBank (accession number AF097635). (b) Developmental Northern blot analysis of the mfl gene. Poly(A)+ RNA was hybridized to a genomic probe, depicted as probe 1 below the map of the region. E, 0–20 h embryos; L1, L2, and L3, first, second, and third instar larvae; P, pupae; F and M, female and male adult flies. The relative amount of RNA loaded in each lane was checked by hybridization with a probe derived from the gene coding for the Drosophila ribosomal protein rp49 (O’Connell and Rosbash, 1984). (c) Hybridization of the same RNA panel shown in b with probe 2 (depicted below the map), specific to mfl maternal mRNA. (d) In situ hybridization of whole mount wild-type ovaries with probe 2. On the left, hybridization with the mfl RNA anti-sense strand (top) and with the sense strand as negative control (bottom); on the right, enlargement of the tip of an ovariole in which the hybridization signal starts to be detected from the early oogenesis stages, within the gerarial region (marked by the arrowhead). (e) Northern blot hybridization of total RNA preparations obtained from 0-2, 2-4, and 4-6 h staged embryos with genomic probe 1. (f) Alignment of MFL and dyskerin amino acid sequences. Black boxed letters highlight identical amino acids, different yet conserved amino acids are on a gray background; block letters on a white background indicate different and nonconserved amino acids. Lines above the sequences indicate putative functional domains; NLS, nuclear localization signal; TruBI and TruBII, regions having homology with bacterial and yeast tRNA pseudouridine synthases; tyr, tyrosine domain; Up, putative uracil-binding pocket. A asterisks on the top indicate the positions of missense mutations so far identified in dyskerin from D KC patients (H. eiss et al., 1998).

**mfl Is an Essential Drosophila Gene Playing a Key Role in rRNA Processing**

When mfl mutants were checked for gene expression, we found that they were all characterized by reduced levels of mfl mRNAs. While the viable, hypomorphic mfl1 allele showed only a modest reduction, mfl expression was strongly disrupted in mfl05 and mfl06 (Fig. 4a), the two alleles causing larval lethality. MFL protein accumulation strictly paralleled the level of mfl mRNAs, so that it was strongly reduced in mfl06 and nearly null in mfl05 (Fig. 4b). Remarkably, the developmental time at which lethality is achieved in these two mutants correlates well with MFL level since, as mentioned above, mfl05 homozygotes die mainly as first instar larvae, while mfl06 animals as second or early third instar larvae (see Fig. 4c). Note that, considering the timing of persistence of maternal rRNA (Winkles et al., 1985), mortality at the first instar is that which...
may be expected for mutations causing severe loss of function of a gene essential for rRNA processing. Taken together, all these data indicated that MFL level may be critical for Drosophila viability. We then attempted to rescue mfl lethal phenotype by ectopically expressing MFL from the heat-inducible hsp70 promoter. mfl05 and mfl06 transgenic animals were then obtained and daily treated at 37°C for 30 min. These heat-shock conditions usually produce amounts of the ectopically expressed protein that largely exceed the wild-type level. However, in our experiments they produced a MFL level just comparable to that present in wild-type flies, even though the induced protein remains quite stable from 6 h to as long as 24 h from the heat-shock pulse (Fig. 4b). Nevertheless, the level of induced MFL is sufficient to allow mfl05 and mfl06 transgenic animals to develop synchronously with their wild-type siblings and show a normal increase in their size. Moreover, 30% of the mfl05 and 80% of the mfl06 transgenic animals reach the pupal stage when grown under daily heat-shock treatment. Moreover, these transgenic animals develop synchronously with their wild-type siblings and show a normal increase in their size.

Figure 4. Molecular and functional characterization of mfl mutants. (a) Northern blot analysis of total RNA extracted from wild-type or mfl animals with a genomic probe including the fourth mfl intron. Female (F) and male (M) adult flies carrying the hypomorphic mfl1 allele or first-instar larvae (L1) carrying the mfl05 and mfl06 alleles were analyzed. (b) Western blot analysis of extracts obtained from wild-type or mfl animals, carrying (+) or not carrying (−) a MFL coding transgene. A affinity-purified rabbit polyclonal anti-MFL antibody, kindly provided by S. Poole, was used. Both wt and mfl homozygous animals were grown under heat shock regimen (30 min/d). As shown, MFL level is reduced in all mfl mutants (lanes −) but reaches nearly the wild-type amount in mfl06 transformed larvae (lanes +) at 6 or 24 h from the heat-shock pulse. (c) Lethal phase (see Materials and Methods) of mfl mutants. In fact, with respect to mfl05 and mfl06, generating equimolar amounts of d and b intermediate molecules. This is consistent with the idea that the mfl mutation primarily affects processing at site 1, generating the type a precursor and by an increase of the d species. This is confirmed by hybridization to probe I (Fig. 5b) which shows that, while in wild-type animals the amount of form b largely exceeds that of d (as expected, being that the b molecule is actively produced by both α and β pathways), in mfl mutants these two forms are detected in similar amounts (Fig. 5d). However, since the processing of form d is inhibited, this species accumulates progressively along larval development (Fig. 5d). Conversely, hybridization to probe III indicated that mfl genetic depletion does not impair site 4
The cleavage of type b molecule, since the amount of form c observed in the mutants exceeds even that of the control (Fig 5 e). We concluded that form c is generated properly, but its further processing is inhibited by mfl mutations. In mfl transgenic flies, MFL over-expression leads to a reversal of all of the effects observed, although the efficiency of pre-rRNA processing is not fully restored. In heat-shocked transformed animals, in fact, MFL expression causes a decrease of pre-rRNA accumulation and an increase in the production of the type a molecule (Fig. 5 c). Processing of the type a precursor also occurs properly, since, as depicted in Fig. 5 d, these larvae show an excess in form b versus form d, although the amount of the b molecule does not reach that observed in wild-type animals. Finally, the amount of form c appears reduced after the heat-shock (Fig. 5 e), indicating that its processing is at least partially restored.

In yeast, lack of Cbf5 gene activity affects not only rRNA processing, but also rRNA pseudouridylation. Thus, we checked the level of modification in wild-type and mfl mutants at several 28S and 18S Ψ sites. With this aim, we used oligonucleotide primers complementary to selected 28S or 18S regions to perform primer extension analyses on CMC-treated Drosophila rRNA. CMC blocks reverse transcription, resulting in a gel band terminating in one residue 3’ of the Ψ site (Bakin and Ofengand, 1993). In planning these experiments, we took advantage of the location of Drosophila 28S rRNA pseudouridines recently reported by Ofengand and Bakin (1997). Instead, none of the 18S Ψ sites checked in our experiments was previously known. In spite of the persistence of maternal rRNA, pseudouridylation appears reduced in mfl larvae at several 28S sites, such as the U2442, U2444, and U2499 residues (Fig. 5 f). Similar reduction was observed at various 18S rRNA sites, such as U830/U831, U840, and U885 (Fig. 5 g), indicating that, as Cbf5, mfl is required for efficient rRNA pseudouridylation.

Minifly Hosts an Intron-encoded Box H/ACA snoRNA

An unexpected feature of the mfl gene structure was revealed by the finding that a small RNA species, ~0.1 kb in length, hybridized specifically with the genomic sequences of the fourth mfl intron, while it was not detected by any cdNA probe. This small RNA was detected in total RNA preparations from all developmental stages and was specifically enriched in the poly(A) - RNA fraction. The length of the small RNA species was accurately determined on denaturing 6% polyacrylamide gels and its 5’ end precisely mapped by primer extension analysis of total
larval RNA using two different oligonucleotides (Fig. 6, b and c). These experiments pointed out that this transcript was ~140 nt long and derived from position +37 to about +176 of the 235-bp-long fourth mfl intron (Fig. 6 a). Since a large number of small nucleolar RNAs are intron encoded (Smith and Steitz, 1997), we checked for the presence of snoRNA elements within the 0.14-kb RNA sequence. Two H boxes (consensus ANANNA) and a 3′ terminal ACA element were found (Fig. 6 a); in addition, the predicted secondary structure of the mfl intron-encoded RNA (Fig. 6 d) conformed well to the hairpin-hinge-hairpin-tail architecture common to most yeast and vertebrate box H/ACA snoRNAs (Ganot et al., 1997b). Two short regions of complementarity between the mfl intron-encoded RNA and Drosophila 18S rRNA were also found (Fig. 6 e). As noticed by Ganot et al. (1997a), short regions of pairing with rRNA flank the site of pseudouridylation, allowing the positioning of the residue to be isomerized at the base of the stem, at the first unpaired position before the 3′ snoRNA helical segment. The pseudouridine selected is found to be separated from the H or ACA box by 14 or, in a few cases, by 15 nucleotides. On the basis of these observations, the rRNA pairing properties of the mfl intron-encoded RNA predicted it may direct pseudouridylation of Drosophila 18S rRNA at position U1820 (Fig. 6 f). Primer extension analysis on CMC-alkali treated (lanes 2 and 3) rRNAs were analyzed by primer extension using an snoRNA complementary to the selected Drosophila 18S rRNA sequences (see Materials and Methods). In lanes 1 and 2, RNA was extracted from wild-type larvae; on lane 3, from mfl05 larvae. Lanes A, G, C, and T are dideoxy sequence reactions performed by using the same primer and a plasmid carrying Drosophila 18S rRNA.
rRNA preparations obtained from mfl05 first instar larvae, we found that pseudouridylation was reduced not only at U 1820, but also at U 1821 and U 1822 residues (Fig. 6 f, lane 3). This result may be explained by the widespread inhibition of rRNA pseudouridylation observed in mfl mutants. Further experiments are thus required to define the specific functional role, if any, played by the mfl intron-encoded RNA.

Finally, we checked the localization of the mfl intron-encoded RNA by in situ hybridization experiments to whole mount ovariety preparations. This analysis showed that a 0.14-kb RNA-specific antisense probe exclusively labeled the nucleoli (Fig. 7, a and b) as it occurs in each tested embryonic or larval tissue (not shown). Specific nucleolar localization may also be observed for MFL (Fig. 7 c), whose ubiquitous expression resulted from both immunolocalization data and histochemical staining of lacZ activity in mfl flies (data not shown). In ovarian tissue preparations we noticed that the protein occasionally diffuses into the cytoplasm in several patches of follicle cells. A s judged by the presence of well defined, round-shaped nucleoli having morphologically well distinguishable nucleoli (Fig. 7 c), these cells should not be in or around mitosis. Moreover, cytoplasmic diffusion can be observed also after stage 10b of oogenesis, when follicular cells endocycles are reported to be terminated (Calvi et al., 1998). It is thus plausible that occasional MFL cytoplasmic localization may be related to ability to carry out nucleolus-cytoplasmic shuttling, as proposed for NA P57 in rat cells (Meier and Blöbel, 1994).

Taken together, the experiments reported indicate that mfl hosts, in its fourth intron, a box H/ACA snoRNA gene, the first member of this class to be identified so far in Drosophila. We have called this gene snoH1 and suggest that it is functionally equivalent to the human U70 snoRNA gene.

Discussion

We reported the cloning of the D. melanogaster mfl gene and established that it encodes an ubiquitous nucleolar protein essential for Drosophila viability and female fertility. Our data also showed that mfl is closely related to the other members of the Cbf5p family so far characterized from higher eukaryotes, the rat Nap57 and the human gene responsible for the X-linked dyskeratosis congenita disease. A s cogently predicted (Luzzatto and Kradimtzis, 1998), flies carrying mutations in the Drosophila DKC1 orthologue show a pleiotropic phenotype very similar to that caused by mutations that affect the synthesis of ribosomal RNA. In fact, we found that mfl loss-of-function mutations impair rRNA processing and lead to accumulation of rRNA precursors. Although these effects are very similar to those caused by Cbf5 genetic depletion, yeast mutations preferentially affect the production of mature 18S rRNA (Lafontaine et al., 1998), while mfl mutations cause similar reduction of 18S and 28S rRNA species. It would be of interest to know whether this is due to a distinctive feature of Drosophila rRNA processing pathways, or whether it reflects a general property of rRNA processing in higher eukaryotes.

In addition to affecting rRNA maturation, mfl loss-of-function causes reduced levels of pseudouridylation at several 28S and 18S rRNA sites, suggesting that gene activity might be required for fully efficient rRNA pseudouridylation. A gain, these results are reminiscent of those obtained in yeast (Lafontaine et al., 1998), and outline the existence of a link between rRNA processing and rRNA pseudouridylation in eukaryotes. By mapping the protein domains conserved among members of the Cbf5p family and investigating the definition of their functional roles, significant information should be generated about the functional role played by rRNA pseudouridylation, which still remains elusive. Although pseudouridylation of eukaryotic rRNA is not observed predominantly on the primary rRNA transcripts before nucleolytic processing, this type of modification is not required for efficient processing of 25S yeast rRNA (Bousquet-Antonelli et al., 1997). It has been suggested that pseudouridylation can contribute to rRNA folding, rRNPs assembly, and ribosomal subunit assembly (Lane et al., 1995; Maden, 1990; Ofengand et al., 1995). Other hypotheses, such as subtle enhancing of ribosomal functions or influencing fidelity of codon recognition, have also been proposed (Ofengand and Bakin, 1997).

A n additional role that could be suggested for MFL is based on the observation that it can occasionally diffuse within the cytoplasm. A s previously suggested for NA P57 in rat cells, it is tempting to speculate that this may possibly reflect the ability of MFL to structure and export pre-ribosomal RNP particles into the cytoplasm. If confirmed, this would strongly support the view that members of this family are multifunctional proteins involved in different aspects of ribosome biogenesis. It is possible that these
proteins may constitute essential components of a single multifunctional complex or, alternatively, they represent common components of structurally and functionally different RNP particles. The definition of the functional interactions required to carry out such a variety of functions will help to clarify this point.

Remarkably, the identification and the characterization of mutations disrupting mfl gene expression has led to establishing the first animal model system for the study of the X-linked dyskeratosis congenita human disease. Some of the results reported here may immediately provide useful information for the comprehension of the molecular basis of the DKC disease. A first relevant point concerns the observation that none of the mfl mutations so far isolated disrupts the gene coding region. Thus, each Drosophila mutant line has certainly quantitative and not qualitative alterations of the gene product which causes the pleiotropic abnormalities observed. The level of MFL protein was found to be critical, and a simple dose-effect rule may be derived: when the protein level is below a crucial threshold, mortality ensues. Instead, while the protein level is lowered but still stands above a critical threshold, the viable, hypomorphic mfl<sup>1</sup> phenotype is reached. By analogy, it can be suggested that in man the level of dyskerin activity may be one of the critical parameters able to trigger the DKC disease. The finding that DKC mutations mapped so far all affect the dyskerin coding region (Hess et al., 1998) is in only apparent contrast with that found in Drosophila. In fact, it is reasonable to suppose that, as observed in Drosophila, total or severe loss-of-function mutations should not be compatible with life. Mutations recovered in patients might be those causing partial loss-of-function, so that the level of dyskerin activity is still compatible with survival. Accordingly, DKC patients might carry hypomorphic mutations, the human counterparts of the viable mfl<sup>1</sup> phenotype. Whether these hypomorphic phenotypes are simply a consequence of the inadequate mature rRNA level or are, at least partially, caused by abnormal accumulation of intermediate rRNA species is an important point which deserves further investigation. A further issue concerns the observation that, although MFL and dyskerin are ubiquitous proteins, phenotypic abnormalities are, in Drosophila as in man, restricted to only certain tissues. Since the gene product is presumed to be critically important for protein synthesis in every cell of the body, the finding that abnormalities are developed only by selected cell types is quite surprising. However, if it is accepted that the level of protein activity may be a critical parameter, then it is reasonable to suppose that the amount of properly processed rRNA may be sufficient in cells having a slow growth rate, while in highly proliferating tissues or in cells sustaining a high rate of protein synthesis this would not the case, and degenerative cell defects could progressively be accumulated. Interestingly, inhibition of protein synthesis is known to be one of the stimuli capable of inducing apoptotic cell death, probably by decreasing the levels of essential proteins or by inhibiting the synthesis of proteins that normally suppress the spontaneous activation of apoptosis (reviewed by Wertz and Hanley, 1996). In mfl<sup>1</sup> ovaries, one of the Drosophila tissues where morphological abnormalities can be observed, we found that degeneration is specifically accompanied by apoptotic cell death. This observation might also suggest a role for apoptosis in the progressive clinical manifestations of the DKC disease.

Finally, we showed that mfl gene organization is intriguing, and leads to the identification of the first member of the box H/ACA class of snoRNA<sub>sa</sub> described so far in Drosophila. As in the case of the snoH<sub>1</sub> gene described here, most of the snoRNA<sub>sa</sub>s are intron encoded, and snoRNA host genes often encode proteins involved in translation or ribosome biogenesis (reviewed by Smith and Steitz, 1997). These intron-encoded snoRNA<sub>sa</sub> are co-transcribed with their host pre-mRNA and their accumulation is splicing-dependent, since they are released from the excised intron by exonucleolytic processing. Our observation that snoH<sub>1</sub> RNA and mfl mRNA levels are reduced in parallel in each mfl mutant line strongly suggests that snoH<sub>1</sub> RNA processing is linked to the splicing of the mfl primary transcript. This feature, which allows coordinated regulation of the host protein and the intron encoded snoRNA<sub>s</sub>, may hinder a precise definition of the specific functional role played by each product. With regard to snoH<sub>1</sub>, it cannot be excluded, in principle, that it may be required for Drosophila viability and that its depletion might contribute to the generation of mfl phenotype. However, we have observed that snoH<sub>1</sub> has little, if any, effect on mfl phenotypic rescue when over-expressed in mfl transgenic flies, either in the presence or in the absence of MFL overexpression (Giordano, E., and M. Furia, unpublished data). This is not surprising, given that all box H/ACA snoRNA<sub>s</sub> found in yeast, with the exception of snr30 (Morrisey and Tollerv, 1993) and snr10 (Tollever, 1987; Tollever and Guthrie, 1985), are dispensable for viability. It will now be interesting to determine whether this type of gene organization is restricted to mfl or is shared by other members of this conserved gene family.

We are grateful to S. Poole for the generous gift of the affinity-purified anti-MFL antibody and to D. Bopp and T. Laverty for sending Drosophila stocks. We thank J. Guariglia, L. Luzzatto, S. Banfi, and P. Delli Bovi for encouragement during the course of this work, R. Calogero and G. Iazzetti for friendly and helpful computational assistance, and G. Falcone for excellent photographic assistance. S. Senger was supported by a Ph.D. fellowship from European Union. This work was supported by funds of MURST PRIN Project Molecular regulation of D evolution to M. Furia. Received for publication 12 November 1998 and in revised form 16 February 1999.

References


of eukaryotic ribosomal RNA. EMBO (Eur. Mol. Biol. Organ.) J. 16:4770

747


