Acta Biochimica Polonica

Vol. 46 No. 2/1999

377-389

QUARTERLY

This paper is dedicated to the memory of Professor Bronisław Filipowicz

Review

Structure and biogenesis of small nucleolar RNAs acting as guides for ribosomal RNA modification

Witold Filipowicz✉, Pawel Pelczar, Vanda Pogacic and François Dragon

Friedrich Miescher-Institut, 4002 Basel, Switzerland

Received: 17 March, 1999

Keywords: nucleolus, RNA modification, rRNA processing, snoRNAs, snoRNPs, modified nucleotides

Maturation of pre-ribosomal RNA (pre-rRNA) in eukaryotic cells takes place in the nucleolus and involves a large number of cleavage events, which frequently follow alternative pathways. In addition, rRNAs are extensively modified, with the methylation of the 2'-hydroxyl group of sugar residues and conversion of uridines to pseudouridines being the most frequent modifications. Both cleavage and modification reactions of pre-rRNAs are assisted by a variety of small nucleolar RNAs (snoRNAs), which function in the form of ribonucleoprotein particles (snoRNPs). The majority of snoRNAs acts as guides directing site-specific 2'-*O*-ribose methylation or pseudouridine formation. Over one hundred RNAs of this type have been identified to date in vertebrates and the yeast *Saccharomyces cerevisiae*. This number is readily explained by the findings that one snoRNA acts as a guide usually for one or at most two modifications, and human rRNAs contain 91 pseudouridines and 106 2'-O-methyl residues. In this article we review information about the biogenesis, structure and function of guide snoRNAs.

Processing and packaging of ribosomal RNAs (rRNAs) into ribosomal particles in eukaryotic cells takes place in the nucleolus. Ribosomal RNAs are synthesised as long 35/47S precursors (pre-rRNAs) which are processed into mature 18S, 5.8S and 25/28S rRNAs. The maturation of rRNAs is an extremely complex process. It occurs concomitantly with packaging of RNA into ribonucleoprotein structures containing most of ribo-

✉Address for correspondence: W. Filipowicz; Friedrich Miescher-Institut, P.O. Box 2543, CH-4002 Basel, Switzerland; phone: (xx 61) 697 4128; fax: (xx 61) 697 3976; e-mail: Filipowi@fmi.ch

Abbreviations: snoRNA, small nucleolar RNA; RNP, ribonucleoprotein; rRNA, ribosomal RNA; pre-rRNA, precursor ribosomal RNA.

somal proteins and also nucleolar proteins associating only transiently with the nascent ribosomes. The precursor rRNA is processed into mature rRNAs through a series of nucleolytic cleavages. A large number of RNA intermediates is produced and processing events may follow alternative pathways. In addition to nucleolytic RNA processing, nucleotides in pre-rRNAs undergo two major types of modifications: methylation of the 2′-hydroxyl group of sugar residues (2′-O-methylation) and conversion of uridines to pseudouridines (Ψ) (pseudouridylation) (for review, see Maden, 1990; Maxwell & Fournier, 1995; Venema & Tollervey, 1995; Sollner-Webb *et al.*, 1996).

Maturation and modification of pre-rRNAs is assisted by a large number of small nucleolar RNAs (snoRNAs), which function in the form of ribonucleoprotein particles (snoRNPs). Some snoRNAs, such as vertebrate U3, U8 and U14, or RNase MRP in yeast, are required for endonucleolytic cleavages of the 35/47S pre-rRNA or resulting intermediates (Filipowicz & Kiss, 1993; Maxwell & Fournier, 1995; Morrissey & Tollervey, 1995; Venema & Tollervey, 1995; Sollner-Webb *et al.*, 1996). However, the majority of snoRNAs acts as guides specifying site-specific 2′-*O*-ribose methylation or Ψ formation (Cavaillé *et al.*, 1996; Kiss-László *et al.*, 1996; Ganot *et al.*, 1997a; 1997b; Ni *et al.*, 1997). Since human rRNAs contain 91 pseudouridines and 106 2′-O-methyl residues (in yeast, these numbers are 43 and 55, respectively) (Maden, 1990; Tollervey & Kiss, 1997; Ofengand & Bakin, 1997; Lafontaine & Tollervey, 1998), the number of guide snoRNAs required for specifying these modifications must be very large. Consistant with this, one hundred and fifty RNAs of this type have been identified to date in vertebrates and yeast (Tollervey & Kiss, 1997; Lafontaine & Tollervey, 1998; Qu *et al.*, 1999; and references therein). In this article we review information about the structure and function of guide snoRNAs. We also discuss mechanisms of their biogenesis, with particular emphasis on processing of intron-encoded snoRNAs.

STRUCTURE AND FUNCTION OF GUIDE snoRNAs

All known vertebrate and yeast snoRNAs fall into two major classes, based on conserved sequence motifs (Balakin *et al.*, 1996; Tollervey & Kiss, 1997; Lafontaine & Tollervey, 1998) (Fig. 1). One class is referred to as C/D family, because its members contain short sequence elements RUGAUGA (box C) and CUGA (box D), located near their 5′ and 3′ ends, respectively. The C and D boxes are usually flanked by short inverted repeats. Many box C/D snoRNAs contain additional C-like and D-like elements referred to as boxes C′ and D′ (Fig. 2). The box C/D snoRNAs are associated with the highly conserved nucleolar protein fibrillarin (Maxwell & Fournier, 1995; Tollervey & Kiss, 1997; Lafontaine & Tollervey, 1998).

Members of the second class of snoRNAs, known as H/ACA, have a characteristic structure consisting of two large stem-loops. These stem-loops are separated by a single-stranded sequence containing a conserved hinge element (box H) with a consensus ANANNA. Another conserved sequence, a trinucleotide ACA, is present downstream of the 3′-terminal stem-loop (Fig. 2). The best characterized protein associated with the H/ACA snoRNAs in yeast is an essential protein Gar1p (Balakin *et al.*, 1996; Ganot *et al.*, 1997b; Bousquet-Antonelli *et al.*, 1997). Boxes C/D and H/ACA are required for formation of respective snoRNPs and for nucleolar accumulation of snoRNAs.

The MRP RNA does not belong to either of the two families discussed above. This RNA is a component of RNase MRP, a ribonucleoprotein enzyme related in structure to RNase P. In yeast, RNase MRP is involved in the endonucleolytic cleavage of pre-rRNA in a region upstream from 5.8S rRNA; following this

	Vertebrates	Yeast
Box C/D:		
	U3, **U8**, **U13**, **U14**, U15, U16, U18, U20, U21, **U22**, U24, U25, U26, U27, U28, U29, U30, U31, U32, U33, U34, U35, U36, U37, U38, U39, U40, U41, U42, U43, U44, U45, U46, U47, U48, U49, U50, U51, U52, U53, U54, U55, U56, U57, U58, U59, U60, U61, U62, U63, U73, U74, U75, U76, U77, U78, U79, U80, U81	**U3**, **U14**, U18, U24, snR4, snR13, snR38, snR39, snR39b, snR40, snR41 snR45, snR47, snR48, snR50, snR51, snR52, snR53, snR54, snR55, snR56, snR57, snR58, snR60, snR61, snR62, snR63, snR64, snR65, snR66, snR67, snR68, snR69, snR70, snR71, snR72, snR73, snR74, snR75, snR76, snR77, snR78, snR190
Box H/ACA:		
	U17 (E1), E2, **E3**, U19, U23, U64, U65, U66, U67, U68, U69, U70, U71, U72	snR3, snR5, snR8, snR9, **snR10**, snR11 **snR30**, snR31, snR32, snR33, snR34, snR35, snR36, snR37, snR42, snR43, snR44, snR46, snR49, snR189
Others:		
	7-2/MRP	**MRP**

Figure 1. Classification of snoRNAs.

The list of snoRNAs (grouped as "Box C/D", "Box H/ACA", and "Others") is given for vertebrates and yeast. SnoRNAs involved in processing reactions are indicated in bold.

cleavage, the mature 5′-end of 5S rRNA is produced by exonucleolytic processing (reviewed in Morrissey & Tollervey, 1995; Tollervey & Kiss, 1997). The RNase MRP particle contains nine different proteins, eight of them shared with RNase P (Chamberlain *et al.*, 1998).

Only very few of the C/D or H/ACA snoRNAs are involved in cleavage reactions. The box C/D U3, U13, U14 and U22 RNAs are required for 18S rRNA production, and U8 is required for processing of 5.8S and 28S rRNAs. The box H/ACA snoRNAs snR10 and snR30 in yeast, and U17 and E3 in mammals, participate in 18S rRNA processing but their exact role is not understood (reviewed in Venema & Tollervey, 1995; Enright *et al.*, 1996; Sollner-Webb *et al.*, 1996). Apart from RNase MRP, none of the snoRNPs has been demonstrated to possess nuclease activity.

The vast majority of the known C/D and H/ACA snoRNAs act as guides for site-specific 2′-*O*-ribose methylation (C/D RNAs) or pseudouridylation (H/ACA RNAs). The box C/D guide snoRNAs generally contain at least 10-nt-long sequence that is complementary to a region in rRNA that is a target of methylation. Mismatches within this complementary region strongly inhibit 2′-*O*-ribose methylation. The sequence complementary to rRNA is always located immediately upstream of the box D or D′, and the position in rRNA which is complementary to the fifth nucleotide from box D (or D′) undergoes modification (Fig. 2). SnoRNAs containing sequences complementary to rRNA upstream of

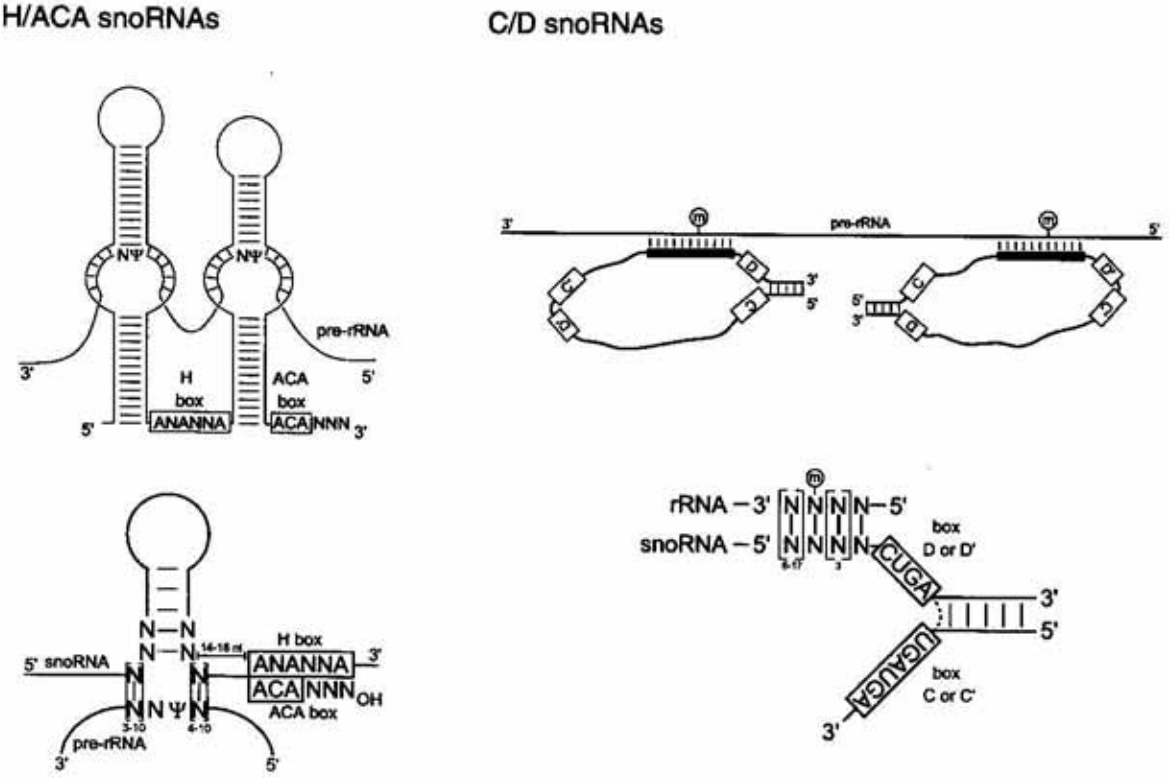

Figure 2. Structure of snoRNAs and their interaction with rRNAs.

The schematic secondary structures of the H/ACA and C/D class snoRNAs. Regions of snoRNAs, which are complementary to pre-rRNA are indicated and the conserved sequences (elements H and ACA, and C/C' and D/D') are boxed. Models for the selection of specific positions to be modified are shown in lower panels. In the case of the H/ACA box snoRNAs, formation of Ψ in rRNA occurs on the unpaired U residue positioned between two regions of snoRNA-rRNA complementarity. This U residue is usually at a distance of 14 to 16 nucleotides from box H or ACA. With the C/D box snoRNAs, 2'-*O*-ribose methylation (m) is performed on the rRNA residue that is base-paired to the fifth position upstream from box D (or D').

both the D and D' boxes specify two different sites of methylation. In a case of some snoRNAs (vertebrate U32 and U36; Nicoloso *et al.*, 1996), one of the targets is in 18S rRNA while another in 28S rRNA. It is not yet known whether methylase activity is an integral component of snoRNPs (Cavaillé *et al.*, 1996; Kiss-László *et al.*, 1996; Tycowski *et al.*, 1996b; Bachellerie & Cavaillé, 1997; Smith & Steitz, 1997; Tollervey & Kiss, 1997).

In box H/ACA snoRNAs, short (3–10 nt) sequences complementary to rRNA are present in the single-stranded internal loops interrupting either one or both stem-loops characteristic of this class of snoRNAs. The U residue to undergo isomerisation is not base-paired but is flanked by two short helices formed by complementary interactions of snoRNA and rRNA. The position of pseudouridine is always 14–16 nucleotides from the box H or ACA (Ganot *et al.*, 1997a; 1997b; Ni *et al.*, 1997; Tollervey & Kiss, 1997) (Fig. 2).

Guide snoRNAs, similarly as snoRNAs involved in pre-rRNA cleavage reactions or spliceosomal U snRNAs, are associated with proteins in snoRNPs. Three proteins common to all investigated C/D box snoRNAs have been identified. One of them is a highly evolutionarily conserved protein known as fibrillarin in vertebrates and Nop1p in yeast. Fibrillarin/Nop1p contains a domain rich in glycines and arginines (GAR domain) and an RBD-type RNA recognition motif. Nop1p is essen-

tial for growth in yeast (Maxwell & Fournier, 1995; Tollervey & Kiss, 1997; Lafontaine & Tollervey, 1998). Two other proteins are Nop56p and Nop58p. These proteins are structurally related (45% identity) and are essential for growth. Nop58p but not Nop56p is required for accumulation of all tested C/D class snoRNAs in yeast (Lafontaine and Tollervey, personal communication).

The H/ACA box snoRNAs in yeast are associated with four essential proteins: Gar1p, Cbf5, Nhp2p and Nop10p (Bousquet-Antonelli *et al.*, 1997; Henras *et al.*, 1998; Lafontaine & Tollervey, 1998; Lafontaine *et al.*, 1998; Watkins *et al.*, 1998). Gar1p, the best characterized protein, is highly conserved in evolution and consists of the central core and two terminal GAR domains (reviewed by Maxwell & Fournier, 1995; Bousquet-Antonelli *et al.*, 1997; Tollervey & Kiss, 1997; Lafontaine & Tollervey, 1998). Cbf5p is strongly homologous to the *Escherichia coli* tRNA:Ψ55 pseudouridine synthase (Lafontaine *et al.*, 1998; Watkins *et al.*, 1998). It is very likely that Cbf5p, and its mammalian counterpart NAP57, are enzymatic components of snoRNPs responsible for pseudouridylation of rRNA (Meier & Blöbel, 1994; Lafontaine *et al.*, 1998; Watkins *et al.*, 1998). Genetic depletion of any of the four proteins associated with the H/ACA snoRNAs results in inhibition of rRNA pseudouridylation and pre-rRNA processing. Depletion of Cbf5p, Nhp2p and Nop10p also prevents accumulation of all tested H/ACA box snoRNAs, suggesting that these proteins are core components of RNPs. Gar1p is not required for accumulation of snoRNAs but is required for the function of the H/ACA snoRNPs (Bousquet-Antonelli *et al.*, 1997; Henras *et al.*, 1998; Lafontaine & Tollervey, 1998; Lafontaine *et al.*, 1998; Watkins *et al.*, 1998). Electron microscopy of purifed yeast H/ACA box snoRNPs, snR30 and snR42, revealed that these two RNPs have a similar bipartite structure. The two domains in the particle may represent a set of core proteins bound to each of the two stems characteristic of all H/ACA class snoRNAs (Watkins *et al.*, 1998).

STRATEGIES OF snoRNA BIOSYNTHESIS AND PROCESSING OF snoRNAs FROM INTRONS

Eukaryotic cells use three different strategies to synthesise snoRNAs (Fig. 3). Some snoRNAs, such as U3, U8, 7-2/MRP, and many yeast snoRNAs, are individually transcribed from independent genes. Another strategy of expression, described for plants and the yeast *Saccharomyces cerevisiae*, involves processing from polycistronic transcripts containing as many as seven different snoRNAs. Processing of polycistronic transcripts into mature snoRNAs in yeast is catalysed by Rnt1p (the yeast ortholog of bacterial RNase III) and different 5′ → 3′ and 3′ → 5′ exonucleases, which also function in the processing of intron-encoded snoRNAs (see below). In yeast, most of the snoRNAs are transcribed from independent genes or are excised from polycistronic transcripts; only very few snoRNAs are encoded in introns of pol II transcription units. In vertebrates, the most common strategy is to process snoRNAs from the excised and linearized introns (reviewed in Leader *et al.*, 1997; Tollervey & Kiss, 1997; Brown & Shaw, 1998; Chanfreau *et al.*, 1998a; 1998b; Pelczar & Filipowicz, 1998; Smith & Steitz, 1998; Qu *et al.*, 1999).

All known mammalian H/ACA snoRNAs and most of the box C/D snoRNAs are encoded in introns of pol II-transcribed genes (Leverette *et al.*, 1992; Maxwell & Fournier, 1995). Using human U17 and U19 H/ACA-type snoRNAs as models, we have previously established a general mechanism for processing of snoRNAs from introns. 5′ → 3′ and 3′ → 5′ exonucleases appear to be involved in the trimming of both ends, and excised and debranched introns act as processing substrates (Kiss & Filipowicz, 1995; Kiss *et al.*, 1996) (Fig. 4). The exonucleolytic trimming of

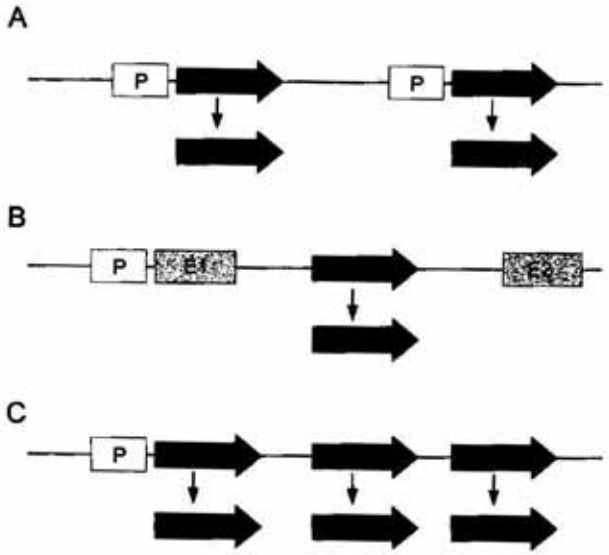

Figure 3. Expression of snoRNAs from independent transcription units (A), from introns of RNA polymerase II-transcribed genes (B), and from polycistronic snoRNA transcripts (C).

The promoter regions (P) are shown as white boxes, exons (E) as gray boxes, and snoRNAs as black arrows.

introns stops at the borders of the snoRNA, most likely due to the snoRNA region being packaged into an RNP. Several lines of experimental evidence support these conclusions. In HeLa cell extracts, processing of snoRNAs from longer intron-like precursors requires free RNA termini; capping of the RNA or its circularization prevent maturation of ends. The most compelling *in vivo* evidence is the demonstration that processing is independent of sequences present upstream and downstream of the snoRNA in the intron, and that only a single snoRNA sequence can be productively processed from one intron: placement of two snoRNAs in tandem in the intron results in accumulation of dimeric snoRNAs (Kiss & Filipowicz, 1995).

Experiments performed with injected *Xenopus* oocytes (Cecconi *et al.*, 1995) and with yeast (Ooi *et al.*, 1998; Petfalski *et al.*, 1998; Villa *et al.*, 1998) also provided evidence for the exonucleolytic mechanism of snoRNA processing from introns. In yeast, genetic inactivation of the intron debranching enzyme, Dbr1p, inhibits snoRNA processing (Ooi *et al.*, 1998; Petfalski *et al.*, 1998). Xrn1p and Rat1p were identified as exonucleases involved in the 5′ → 3′ trimming of intron-encoded snoRNAs (Petfalski *et al.*, 1998). However, the mechanism of snoRNA processing outlined above has some exceptions. Some snoRNAs, such as vertebrate U16 or U18, which are located in poorly spliced introns, are excised from the intron by endonucleases, and only the remaining extra nucleotides flanking the snoRNA are removed exonucleolytically (Caffarelli *et al.*, 1994; 1996). In yeast, U18 snoRNA can be processed from introns by both the exo- and endonucleolytic pathway (Villa *et al.*, 1998).

HOST GENES ENCODING INTRONIC snoRNAs

Characterization of genes that act as hosts for intronic snoRNAs in vertebrates has produced many interesting findings. Most of the host genes analyzed to date encode proteins essential for ribosome biogenesis or function such as ribosomal and nucleolar proteins or translational factors. This observation has led to the speculation that cotranscription of snoRNAs with mRNAs for nucleolar proteins

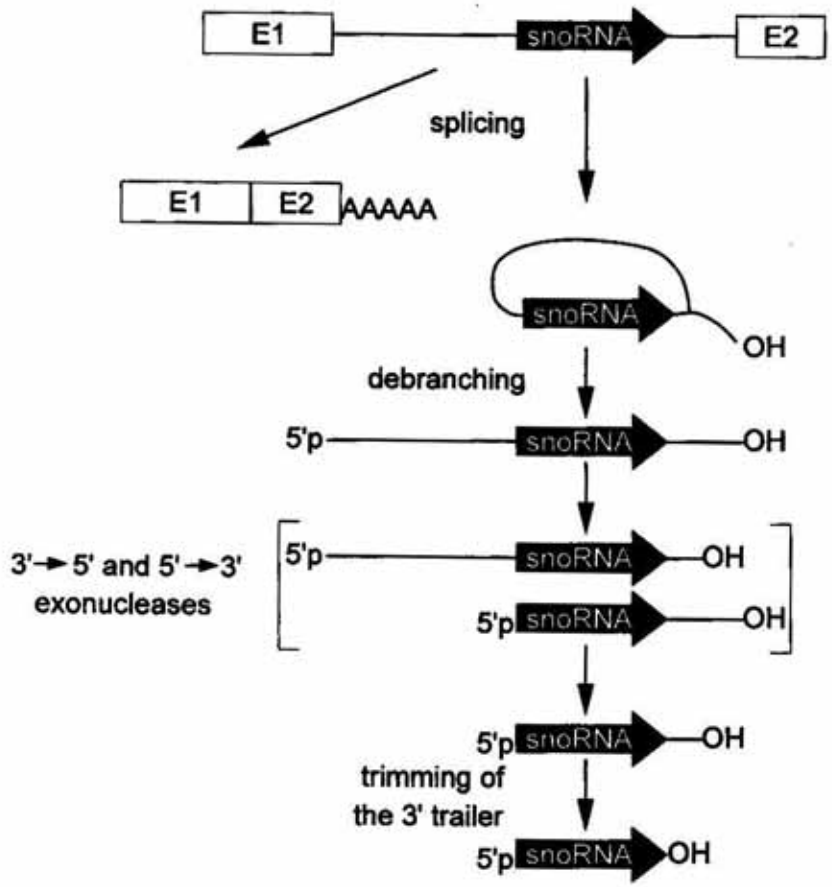

Figure 4. A mechanism for exonucleolytic processing of snoRNAs from excised and debranched introns.

The 5'-phosphate and 3'-OH ends are indicated. The snoRNA sequence (black arrow) associates with snoRNA-specific proteins prior to the final exonucleolytic processing steps.

or translational components may provide a regulatory mechanism to coordinate accumulation of different molecules required for the assembly and function of ribosomes (Sollner-Webb, 1993; Maxwell & Fournier, 1995). Interestingly, Bachellerie and co-workers (Qu *et al.*, 1999) have recently found that promoters of several genes encoding either single unit or polycistronic snoRNAs in yeast contain transcriptional elements similar to those found in ribosomal protein genes, suggesting that expression of both classes of genes may be coregulated.

Not all genes hosting snoRNAs in introns code for proteins. Tycowski *et al.* (1996a; 1996b) have discovered that the gene *UHG*, which harbors intronic snoRNAs U22 and U25-U31, is very unusual. The spliced polyA$^+$ RNAs produced from *UHG* genes in humans, mice, and frogs are not conserved in sequence and have no apparent protein coding potential. We have investigated the organization and expression of the locus encoding intronic snoRNAs U17a and U17b in human and mouse cells (Pelczar & Filipowicz, 1998). In humans, U17 RNAs are transcribed as parts of the three-exon transcription unit, named *U17HG*, positioned approximately 9 kb upstream of the *RCC1* locus. Comparison of the human and mouse *U17HG* genes has revealed that, as in the case of the *UHG* gene, exon sequences are not conserved between the two species and that neither human nor mouse spliced U17HG polyA$^+$ RNAs have a potential to code for proteins (Fig. 5). The finding that, despite its cytoplasmic localization, little if any U17HG RNA is associated with polysomes in HeLa cells, further argues against

an mRNA function of this RNA (Pelczar & Filipowicz, 1998).

Two other non-protein-coding snoRNA host genes have recently been identified. Bortolin & Kiss (1998) have demonstrated that human U19 RNA is encoded in intron 2 of the multiexon *U19HG* gene. Interestingly, the spliced U19HG RNA has a nucleoplasmic rather than cytoplasmic localization. Smith & Steitz (1998) have found that human *gas5* (growth arrest specific transcript 5) gene encodes ten box C/D snoRNAs in its introns. Comparison of human and mouse gas5 genes has revealed no sequence conservation or significant coding potential in spliced exons of either gene.

In summary, four snoRNA host genes which encode polyA$^+$ RNAs having no protein-coding potential have been identified to date. These genes probably act only as vehicles for expression of intron-encoded snoRNAs although it can not be entirely excluded that their spliced polyA$^+$ RNA products also function as regulatory or structural RNAs. RNA folding programs have not identified any obvious secondary structure elements shared by human and mouse UHG, U17HG or gas5 RNAs (Tycowski *et al.*, 1996a; 1996b; Pelczar & Filipowicz, 1998; Smith & Steitz, 1998). Findings that in *UHG*, *U17HG* and *gas5* genes these are snoRNA-encoding introns

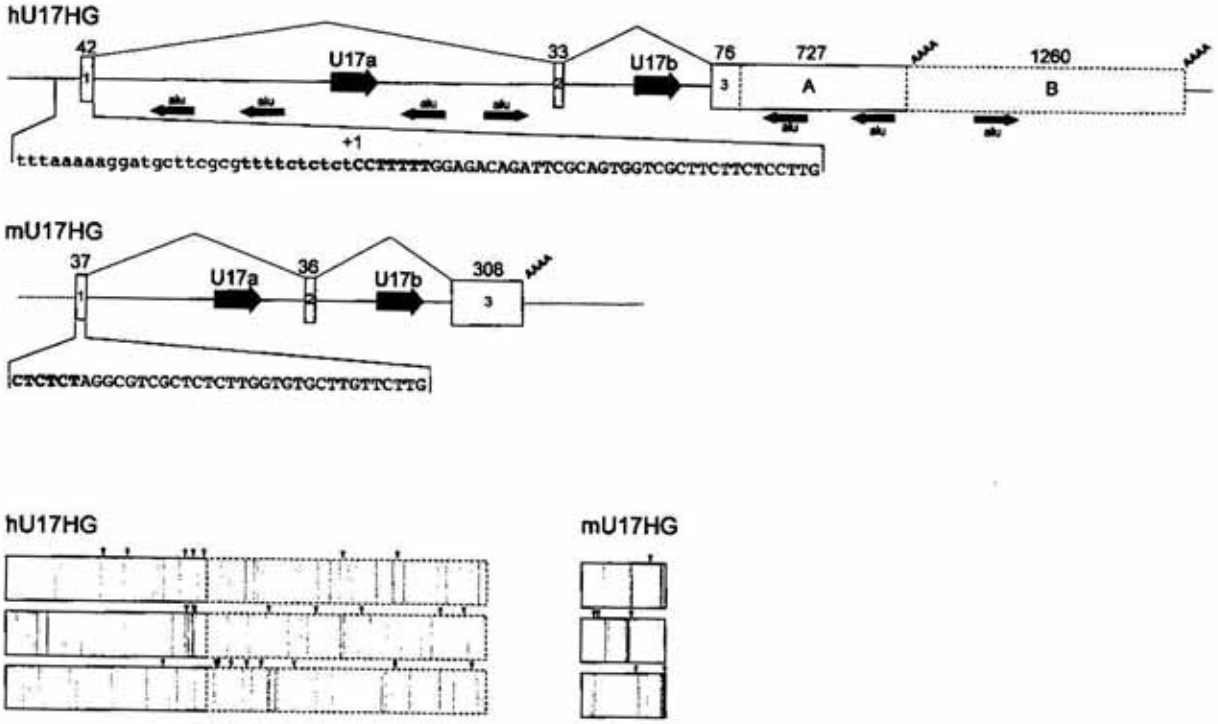

Figure. 5. Schematic structure of human and mouse U17HG genes (upper panels) and analysis of the protein coding potential of spliced U17HG RNAs (lower panels).

Positions of exons and snoRNA regions are denoted by open boxes and black arrows, respectively; sizes and numbers of exons are also indicated. Alu sequences are shown as small black arrows. Human U17HG (hU17HG) RNA is polyadenylated at two different sites (AAAA). Extended version of hU17HG RNA is shown with a dashed line. The transcription-start-site-proximal sequences are shown, with polypyrimidine tracts indicated in bold. (Bottom panels) The coding potential of human (left panel) and mouse (mU17HG; right panel) U17HG RNAs. Positions of AUG codons (black triangles) and stop codons (vertical lines) are indicated for all reading frames of the spliced mouse and human U17HG RNAs. The region corresponding to the extended version of exon 3 in human RNA is drawn with dashed lines. (Reproduced (in modified form), with permission, from Pelczar & Filipowicz (1998)).

and not exons that are evolutionarily conserved and express functional RNAs requires a modification of the current description of exons as the main information-carrying regions of a gene.

Determination of the 5′ terminus of RNAs encoded by *U17HG* and *gas5* genes revealed that transcription of these genes starts with a C residue followed by a polypyrimidine tract (Pelczar & Filipowicz, 1998; Smith & Steitz, 1998). This property makes the genes members of the 5′-terminal oligopyrimidine (5′TOP) family which includes genes encoding ribosomal proteins, some translation factors, and few other house-keeping genes. The polypyrimidine tracts are responsible for upregulation of translation of the 5′TOP mRNAs in response to growth factors or other conditions which require coordinated increased synthesis of proteins making up the translational apparatus (reviewed by Meyuhas *et al.*, 1996; Amaldi & Pierandrei-Amaldi, 1997). Interestingly, other known snoRNA host genes, including non-protein-coding UHG and U19HG genes, have features of the 5′TOP genes (Tycowski *et al.*, 1996a; Bortolin & Kiss, 1998). Similar characteristics of the transcription start site regions in snoRNA host and ribosomal protein genes raise the possibility that expression of components of ribosome biogenesis and translational machineries is coregulated. It is more plausible that coordination of expression of snoRNA hosts and the genes coding for translational components occurs at the transcriptional rather than the translational level. It is rather unlikely that 5′-terminal oligopyrimidine tracts have been conserved in UHG, U17HG and gas5 RNAs in order to regulate translation of these apparently non-protein-coding RNAs (Tycowski *et al.*, 1996a; Pelczar & Filipowicz, 1998; Smith & Steitz, 1998). The oligopyrimidine tracts in 5′TOP genes are usually present not only downstream but also upstream of the start site C residue (Meyuhas *et al.*, 1996; Amaldi & Pierandrei-Amaldi, 1997) and may play a role in transcription of the 5′TOP genes. Perry and co-workers (Hariharan & Perry, 1990; Safrany & Perry, 1995) have reported that integrity of the start site oligopyrimidine tracts in mouse ribosomal protein genes S16 and L30 is important for efficient and precise initiation of their transcription. Although transcriptional role for the oligopyrimidine tract in non-protein-coding host genes is the most probable one, it is possible that these sequences have a different function, more directly related to the processing of snoRNAs from primary transcripts (Pelczar & Filipowicz, 1998; Smith & Steitz, 1998).

ADDITIONAL FUNCTIONS FOR GUIDE snoRNAs AND SOME UNANSWERED QUESTIONS

Two recent reports indicated that guide RNAs specifying nucleotide modifications also function in processes other than rRNA modification. Tycowski *et al.* (1998) have identified in vertebrates two C/D class guide RNAs, named mgU6-47 and mgU6-77, which specify 2′-O-methylation of nucleotides in the spliceosomal U6 snRNA. Similarly as all known C/D box snoRNAs, the two guide RNAs targeting U6 modifications are localized in the nucleolus. The mgU6-77 RNA, in addition to specifying modification of the C residue 77 in U6 snRNA, also directs the 2′-O-methylation of 28S rRNA at position C2970. Since U6 RNA is nucleoplasmic, the latter finding raises an interesting question of where in the cell the 2′-O-methylation of U6 snRNA occurs.

Mitchell *et al.* (1999) have uncovered that the 3′-terminal domain of the RNA component of telomerase in mammals has the H/ACA snoRNA-like structure. Integrity of this domain (e.g., intact H and ACA boxes) is essential for accumulation of the telomerase RNA *in vivo*, and for telomerase activity. It is not known whether the telomerase H/ACA domain guides pseudouridylation of any

RNA target. It may be just required for proper 3′ end processing and stabilization of telomerase RNA by protecting it against digestion by 3′ → 5′ exonucleases. Interestingly however, replacement of the authentic H/ACA domain in the telomerase RNA by the H/ACA snoRNA U64 is sufficient for 3′ end processing and accumulation of the resulting chimeric RNP but not for its activity. Hence, the H/ACA domain in telomerase RNA may also contribute to the assembly and/or function of telomerase.

The specific function of nucleotide modifications introduced in rRNA by guide snoRNPs is not known. The modifications are largely restricted to the universally conserved cores of mature rRNAs, likely to play a fundamental role in ribosome function. However, despite this evolutionarily conserved localization of modifications to functional regions of rRNA, disruption of expression in yeast of as many as seven snoRNAs at a time has no effect on cell viability or growth (reviewed in Maxwell & Fournier, 1995; Bachellerie & Cavaillé, 1997; Lafontaine & Tollervey, 1998; Qu *et al.*, 1999). It is possible that modified nucleotides contribute to more global phenomena rather than act individually and deletion of tens of snoRNA genes would be required in order to see the effect. One can envisage that the modifications are required for rRNA folding or contribute to some of the many interactions, either RNA–RNA or RNA–protein, the ribosome is involved in. Addition of methyl groups generates more hydrophobic surfaces or may stabilize RNA stems by constraining the sugar residues into the more rigid C3′-endo conformation (Kowalak *et al.*, 1994; Lafontaine & Tollervey, 1998). On the other hand, usage of Ψs increases the hydrogen bonding potential (three hydrogen bonds instead of two). It is also possible that blocking of the 2′-hydroxyl groups by methylation may protect crucial regions of RNA from hydrolytic degradation. Additional informations as to possible functions of modified nucleotides in rRNA and evolution of guide snoRNAs can be found in articles by Maden (1990), Ofengand & Bakin (1997), and Lafontaine & Tollervey (1998).

Questions addressing possible origins of non-protein-coding snoRNA host genes are also interesting though difficult to answer conclusively. Several scenarios can be envisaged (Pelczar & Filipowicz, 1998). (I) Exons of these genes originally encoded a protein but this property was lost during evolution. (II) The host exons may have encoded a structural or regulatory RNA, whose function became obsolete with time or which we have not yet identified; (III) The genes originated from polycistronic snoRNA genes, similar to genes expressed in plants and yeast, by conversion of the inter-snoRNA spacers into spliceable exons. The last possibility would imply that units encoding poly-snoRNAs are evolutionarily very old. Such genes could have arisen by duplication of single snoRNA units at a time when the complexity of rRNA modification was increasing. Characterization of snoRNA transcription units in additional distantly related organisms might throw more light on origins of the non-protein-coding snoRNA host genes.

REFERENCES

Amaldi, F. & Pierandrei-Amaldi, P. (1997) TOP genes: A translationally controlled class of genes including those coding for ribosomal proteins. *Prog. Mol. Subcell. Biol.* **18**, 1–17.

Bachellerie, J.-P. & Cavaillé, J. (1997) Guiding ribose methylation of rRNA. *Trends Biochem. Sci.* **22**, 257–261.

Balakin, A.G., Smith, L. & Fournier, M.J. (1996) The RNA world of the nucleolus: Two major families of small RNAs defined by different box elements with related functions. *Cell* **86**, 823–834.

Bortolin, M.-L. & Kiss, T. (1998) Human U19 intron-encoded snoRNA is processed from a long primary transcript that possesses little potential for protein coding. *RNA* **4**, 445–454.

Bousquet-Antonelli, C., Henry, Y., Gélugne, J.-P., Caizergues-Ferrer, M. & Kiss, T. (1997) A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *EMBO J.* **16**, 4770–4776.

Brown, J.W.S. & Shaw, P.J. (1998) Small nucleolar RNAs and pre-rRNA processing in plants. *Plant Cell* **10**, 649–657.

Caffarelli, E., Arese, M., Santoro, B., Fragapane, P. & Bozzoni, I. (1994) *In vitro* study of processing of the intron-encoded U16 small nucleolar RNA in *Xenopus laevis*. *Mol. Cell. Biol.* **14**, 2966–2974.

Caffarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P. & Bozzoni, I. (1996) Processing of the intron-encoded U16 and U18 snoRNAs: The conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *EMBO J.* **15**, 1121–1131.

Cavaillé, J., Nicoloso, M. & Bachellerie, J.-P. (1996) Targeted ribose methylation of RNA *in vivo* directed by tailored antisense RNA guides. *Nature* **383**, 732–735.

Cecconi, F., Mariottini, P. & Amaldi, F. (1995) The Xenopus intron-encoded U17 snoRNA is produced by exonucleolytic processing of its precursor in oocytes. *Nucleic Acids Res.* **23**, 4670–4676.

Chamberlain, J.R., Lee, Y., Lane, W.S. & Engelke, D.R. (1998) Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes Dev.* **12**, 1678–1690.

Chanfreau, G., Rotondo, G., Legrain, P. & Jacquier, A. (1998a) Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *EMBO J.* **17**, 3726–3737.

Chanfreau, G., Legrain, P. & Jacquier, A. (1998b) Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J. Mol. Biol.* **284**, 975–988.

Enright, C.A., Maxwell, E.S., Eliceiri, G.L. & Sollner-Webb, B. (1996) 5'ETS rRNA processing facilitated by four small RNAs: U14, E3, U17, and U3. *RNA* **2**, 1094–1099.

Filipowicz, W. & Kiss, T. (1993) Structure and function of nucleolar snRNPs. *Mol. Biol. Reports* **18**, 149–156.

Ganot, P., Bortolin, M.-L. & Kiss, T. (1997a) Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* **89**, 799–809.

Ganot, P., Caizergues-Ferrer, M. & Kiss, T. (1997b) The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.* **11**, 941–956.

Hariharan, N. & Perry, R.P. (1990) Functional dissection of a mouse ribosomal protein promoter: Significance of the polypyrimidine initiator and an element in the TATA-box region. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1526–1530.

Henras, A., Henry, Y., Bousquet-Antonelli, C., Noillac-Depeyere, J., Gelugne, J.-P. & Caizergues-Ferrer, M. (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J.* **17**, 7078–7090.

Kiss, T. & Filipowicz, W. (1995) Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes Dev.* **9**, 1411–1424.

Kiss, T., Bartolin, M.-L. & Filipowicz, W. (1996) Characterisation of the intron-encoded U19 RNA, a new mammalian small nucleolar RNA that is associated with fibrillarin. *Mol. Cell. Biol.* **16**, 1391–1400.

Kiss-László, Z., Henry, Y., Bachellerie, J.-P., Caizergues-Ferrer, M. & Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs. *Cell* **85**, 1077–1088.

Kowalak, J.A., Dalluge, J.J., McCloskey, J.A. & Stetter, K.O. (1994) The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. *Biochemistry* **33**, 7869–7876.

Lafontaine, D.L.J. & Tollervey, D. (1998) Birth of the snoRNPs: The evolution of the modification guide snoRNAs. *Trends Biochem. Sci.* **23**, 383–388.

Lafontaine, D.L.J., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. & Tollervey,

D. (1998) The box H+ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev.* **12**, 527–537.

Leader, D.J., Clark, G.P., Watters, J., Beven, A.F., Shaw, P.J. & Brown, J.W.S. (1997) Clusters of multiple different small nucleolar RNA genes in plants are expressed as and processed from polycistronic pre-snoRNAs. *EMBO J.* **16**, 5742–5751.

Leverette, R.D., Andrews, M.T. & Maxwell, E.S. (1992) Mouse U14 snRNA is a processed intron of the cognate hsc70 heat shock premessenger RNA. *Cell* **71**, 1215–1221.

Maden, B.E. (1990) The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **39**, 241–303.

Maxwell, E.S. & Fournier, M.J. (1995) The small nucleolar RNAs. *Annu. Rev. Biochem.* **64**, 897–934.

Meier, U.T. & Blöbel, G. (1994) NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. *J. Cell Biol.* **127**, 1505–1514.

Meyuhas, O., Avni, D. & Shama, S. (1996) Translational control of ribosomal protein mRNAs in eukaryotes; in *Translational Control* (Hershey, J.W.B., Mathews, M.B. & Sonenberg, N., eds.) pp. 363–388, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mitchell, J.R., Cheng, J. & Collins, K. (1999) A box H/ACA small nucleolar RNA-like domain in the human telomerase RNA 3′ end. *Mol. Cell. Biol.* **19**, 567–576.

Morrissey, J.P. & Tollervey, D. (1995) Birth of the snoRNPs: The evolution of RNase MRP and the eukaryotic pre-rRNA-processing system. *Trends Biochem. Sci.* **20**, 78–82.

Ni, J., Tien, A.L. & Fournier, M.J. (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* **89**, 565–573.

Nicoloso, M., Qu, L.-H., Michot, B. & Bachellerie, J.-P. (1996) Intron-encoded, antisense small nucleolar RNAs: The characterisation of nine novel species points to their direct role as guides for 2′-*O*-ribose methylation of rRNAs. *J. Mol. Biol.* **260**, 178–195.

Ofengand, J. & Bakin, A. (1997) Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts. *J. Mol. Biol.* **266**, 246–268.

Ooi, S.L., Samarsky, D.A., Fournier, M.J. & Boeke, J.D. (1998) Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariat-debranching enzyme: Intron length effects and activity of a precursor snoRNA. *RNA* **4**, 1096–1110.

Pelczar, P. & Filipowicz, W. (1998) The host gene for intronic U17 small nucleolar RNAs in mammals has no protein-coding potential and is a member of the 5′-terminal oligopyrimidine gene family. *Mol. Cell. Biol.* **18**, 4509–4518.

Petfalski, E., Dandekar, T., Henry, Y. & Tollervey, D. (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.* **18**, 1181–1189.

Qu, L.H., Henras, A., Lu, Y.-J., Zhou, H., Zhou, W.-H., Zhu, Y.-Q., Zhao, J., Henry, Y., Caizergues-Ferrer, M. & Bachellerie, J.-P. (1999) Seven novel methylation guide small nucleolar RNAs are processed from common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.* **19**, 1144–1158.

Safrany, G. & Perry, R.P. (1995) The relative contributions of various transcription factors to the overall promoter strength of the mouse ribosomal protein L30 gene. *Eur. J. Biochem.* **230**, 1066–1072.

Smith, C.M. & Steitz, J.A. (1997) Sno storm in the nucleolus: New roles for myriad small RNPs. *Cell* **89**, 669–672.

Smith, C.M. & Steitz, J.A. (1998) Classification of gas5 as a multi-snoRNA host gene and a member of the 5′terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol. Cell. Biol.* **18**, 6897–6909.

Sollner-Webb, B. (1993) Novel intron-encoded small nucleolar RNAs. *Cell* **75**, 403–405.

Sollner-Webb, B., Tycowski, K.T. & Steitz, J.A. (1996) Ribosomal RNA processing in eukaryo-

tes; in *Ribosomal RNA: Structure, Evolution, Processing and Function in Protein Biosynthesis* (Zimmermann, R.A. & Dahlberg, A.E., eds.) pp. 469–490, CRC Press, Boca Raton, FL.

Tollervey, D. & Kiss, T. (1997) Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.* **9**, 337–342.

Tycowski, K.T., Shu, M.-D. & Steitz, J.A. (1996a) A mammalian gene with introns instead of exons generating stable RNA products. *Nature* **379**, 464–466.

Tycowski, K.T., Smith, C.M., Shu, M.-D. & Steitz, J.A. (1996b) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14480–14485.

Tycowski, K.T., You, Z.-H., Graham, P.J. & Steitz, J.A. (1998) Modification of U6 spliceosomal RNA is guided by another small RNA. *Mol. Cell* **2**, 629–638.

Venema, J. & Tollervey, D. (1995) Processing of pre-ribosomal RNA in *Saccharomyces cerevisiae*. *Yeast* **11**, 1629–1650.

Villa, T. & Ceradini, F., Presutti, C. & Bozzoni, I. (1998) Processing of the intron-encoded U18 small nucleolar RNA in the yeast *Saccharomyces cerevisiae* relies on both exo- and endonucleolytic activities. *Mol. Cell. Biol.* **18**, 3376–3383.

Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. & Luhrmann, R. (1998) Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H box/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA* **4**, 1549–1568.