MAMMALIAN RIBOSOMAL PROTEIN S3a GENES AND INTRON-ENCODED SMALL NUCLEOLAR RNAs U73 AND U82

ANA REBANE

MAMMALIAN RIBOSOMAL PROTEIN S3a GENES AND INTRON-ENCODED SMALL NUCLEOLAR RNAs U73 AND U82

ANA REBANE

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I. Metspalu, A., Rebane, A., Hoth, S., Pooga, M., Stahl, J. and Kruppa, J. (1992) Human ribosomal protein S3a: cloning of the cDNA and primary structure of the protein. *Gene* 119, 313–316.

II. Rebane, A., Tamme, R., Laan, M., Pata, I. and Metspalu, A. (1998) A novel snoRNA (U73) is encoded within the introns of the human and mouse ribosomal protein S3a genes. *Gene* 210, 255–263.

III. Rebane, A. and Metspalu A. (1999) U82, a novel snoRNA identified from the fifth intron of human and mouse nucleolin gene. *Biochim. Biophys. Acta* 1446, 426–430.

LIST OF ABBREVIATIONS AND SYMBOLS

3' ETS	3' external transcribed spacer
5' ETS	5' external transcribed spacer
5' TOP	5' terminal oligopyrimidine tract
5' UTR	5' untranslated region
aa	amino acid(s)
bp	basepair(s)
cDNA	complementary DNA
CpG	dinucleotide 5'-CG-3'
DNA	deoxyribonucleic acid
FISH	fluorescence *in situ* hybridization
eIF-2	eukaryotic initiation factor 2
eIF-3	eukaryotic initiation factor 3
EST	expressed sequenced tag
ITS1	internal transcribed spacer 1
ITS2	internal transcribed spacer 2
kb	kilobasepairs
kDa	kilodalton(s)
mRNA	messenger RNA
nt	nucleotide(s)
NLS	nuclear localization signal
PCR	polymerase chain reaction
pre-rRNA	preribosomal RNA
RNA	ribonucleic acid
rp(s)	ribosomal protein(s)
rRNA	ribosomal RNA
snoRNA	small nucleolar RNA
STS	sequence-tagged sites
tRNA	transport RNA
TSP	transcription start point
ψ	pseudouridine

1. INTRODUCTION

Protein synthesis is performed by the ribosomes in all living cells. In eukaryotic cells, the biogenesis of cytoplasmic ribosomes takes place in the nucleolus. 5S RNA, which is synthesized outside the nucleolus by RNA polymerase III and ribosomal proteins (rps), which are synthesized in the cytoplasm, have to be transported to the nucleolus. The 18S, 5.8S and 25/28S ribosomal RNAs are transcribed by RNA polymerase I in the nucleolus as a common precursor RNA (pre-rRNA) containing long external and internal transcribed spacer sequences. During maturation and assembly with the ~80 rps the pre-rRNA undergoes cleavage, folding and extensive site specific modifications, mostly 2'-O-ribose methylation of riboses and conversion of uridine residues to pseudouridine (reviewed in Eichler and Craig, 1994). Several nucleolar proteins and up to 200 small nucleolar RNAs (snoRNAs) in the form of small nucleolar ribonucleoprotein complexes (snoRNPs) assist ribosome biogenesis. After maturation, ribosomal subunits are exported to the cytoplasm where a few late-assembly proteins are adjoined to them.

To ensure a balanced supply of a large number of ribosome constituents as well as sufficient quantity of nonribosomal proteins required for translation, nucleolar proteins and snoRNAs, their synthesis and transport must be coordinated. In vertebrates, almost all genes encoding rps and proteins that are involved in ribosome production or function, belong to the 5' TOP (5' terminal oligopyrimidine) gene family. The expression of 5' TOP genes has thought to be regulated mainly at translational level (reviewed in Amaldi and Pierandrei-Amaldi, 1997). However, recently it has been noticed that all characterized host genes for intron-encoded snoRNAs are 5' TOP genes. This observation raises the possibility that expression of snoRNAs and genes encoding components of translational machinery is coregulated at transcriptional level.

Although the ribosome constituents are known to be essential for growth and development of organisms, the effects of ribosomal mutations and their role in human disorders remain largely to be elucidated. It has been predicted that the genetic defects in ribosomal components would invariably result in early embryonic death in vertebrates. However, a potential role of ribosomal proteins in human diseases has received considerable attention recently, after the rp *S19* gene was found to be mutated in a subset of Diamond-Blackfan anaemia patients (Draptchinskaia *et al.*, 1999).

The first indispensable step in elucidating the functions of ribosomal proteins and other components involved in ribosome biosynthesis or translation, is the characterization of the genes for these components. Present work outlines cloning and characterization of the cDNA for human rp S3a, detailed structural analysis of the human and mouse S3a functional genes and internally encoded snoRNA U73. In addition, identification of a novel snoRNA (U82) from the fifth intron of the human and mouse nucleolin gene is described.

3   9

2. REVIEW OF LITERATURE

2.1. Mammalian ribosomal proteins

2.1.1. Structure of ribosomal proteins

The primary structure of 80 different mammalian ribosomal proteins has been determined so far (Chan *et al.*, 1996, Wool *et al.*, 1995). The rps are small and mostly basic, in mammalians they consist of 25 to 421 aa and have an average pI of 11.05. The rps are well conserved throughout the evolution. Most mammalian rps have counterparts in yeast with an average identity of 60%. More than half rps can be related to archaebacterial rps and almost one third possess eubacterial homologs (reviewed in Wool *et al.*, 1995).

Common structural peculiarities of rps are clustered sets of basic and aromatic amino acids and repetition of short peptide sequences. Severel rps, including the only acidic ones (P0, P1 and P2) contain also specific clusters of acidic amino acids. Generally, it is assumed that possible functions of basic clusters and repeats are associated with the transport of rps from cytoplasm to the nucleolus where the ribosome synthesis occurs, or with RNA binding (Wool *et al.*, 1995). Using X-ray crystallography and NMR spectroscopy, 15 bacterial rp structures have been determined so far. The mapping of conserved aa residues onto these structures has revealed that basic and aromatic aa are frequently exposed and clustered on protein surface and therefore they function probably as RNA-binding sites (Ramakrishnan and White, 1998). Protein regions that are required for nuclear import are generally designated as nuclear localization signals (NLS). NLSs are predominantly short basic amino acid sequences that possess only limited homology to each other and to either continuous or bipartite motifs (Hicks and Raikhel, 1995). Studies on different proteins have shown that other protein regions could also support the transport. For instance, nuclear import and nucleolar accumulation of human rp S7 depends on both, a minimal NLS and an adjacent basic region (Annilo *et al.*, 1998).

In addition to nucleolar localization and RNA binding, the repetitions of short peptide sequences of rps can be required for protein-protein interactions. For example, the rp S2 amino-terminus is extremely glycine and arginine rich, it contains two RGGF motifs and more than ten RG repeats (Suzuki *et al.*, 1991). A similar RG-rich domain of nucleolin has been shown to be capable to bind to 18 different rps (Bouvet *et al.*, 1998).

Several rps (S27, S29, L37, L37a, S27a and L40) hold zinc finger motifs (Chan *et al.*, 1993, Chan *et al.*, 1995). Proteins containing zinc fingers share the potential to coordinate a zinc ion and bind to DNA and seldom to RNA (Berg, 1990). There is no evidence whether the zinc fingers of rps participate in their binding to rRNA at present or do they reflect possible former functions of these proteins. For instance, the rat S29 and its homologs in several species contain an

10

intact zinc finger, while yeast and *E. coli* homologs contain only partial motifs that are probably not able to coordinate zinc ion (Chan *et al.*, 1993). It has been suggested that ribosomal proteins may originate from proteins that earlier had or still have other functions in cells (see also section 2.1.4.). Some of rps could be recruited from a set of proteins that already had amino acid sequences necessary for binding to nucleic acids (Wool *et al.*, 1995, Wool, 1996).

Numerous rps (L7, L13a, L9, L12, L35, L37a, S2 and S9) contain domains similar to the basic region leucine zipper (bZIP) motif that has a potential to mediate protein dimerization and nucleic acid binding (Wool *et al.*, 1995). For rp L7, it has been confirmed that the bZIP element mediates the formation of L7 homodimers that inhibit the translation of mRNAs through interacting with their cognate sites (Neumann *et al.*, 1995). It has not been determined, whether the domains reminiscent to bZIP of other rps are functional.

Many rps are subjected to post-translational processing. S6, P0, P1 and P2 form a group of rps that are phosphorylated. Phosphorylation of P0, P1 and P2 is necessary for their assembly into 60S (Remacha *et al.*, 1995). S6 is a major ribosomal phosphoprotein and is phosphorylated by different kinases in response to various mitogenes (see also sections 2.1.3. and 2.1.4.) (reviewed in Stewart and Thomas, 1994). Three proteins (S27a, L40 and S30) are unique among rps as they are formed by cleavage of a hybrid protein. S27a and L40 are the carboxyl extensions of ubiquitin (Chan *et al.*, 1995) and S30 is a carboxyl extension of ubiquitin-like protein (Olvera and Wool, 1993). The reason for ubiquitinylation of these three rps is unclear. Mass spectrometric analysis of rat 40S rps has shown that many rps are N-terminally demethionylated and/or acetylated (Louie *et al.*, 1996). Ubiquitous methylation of Lys22 within the zinc-finger motif of L40 and tissue-specific methylation of Lys4 in L29 have been reported. The function of either of these modifications is not known (Williamson *et al.*, 1997).

2.1.2. Organization and structure of ribosomal protein genes

In eukaryotes, the rp genes are scattered over the entire genome and they are all transcribed independently. Recently, using intron-specific STS mapping, 75 of 80 human ribosomal protein genes have been mapped on the chromosomes. Both sex chromosomes and at least 20 of the 22 autosomes have been demonstrated to carry one or more rp genes (Kenmochi *et al.*, 1998). In *Saccharomyces cerevisiae*, 59 of the 78 rp genes have two transcriptionally active copies (Planta, 1997). By contrast, all rps except the human S4 (see also section 2.1.4.) are known to be encoded by a single intron-containing functional gene in vertebrates. In addition to one transcriptionally active gene, there are usually more than a dozen promoterless processed pseudogenes scattered all over the genome (Wool *et al.*, 1995). It has been suggested that the processed pseudogenes originate from integration of reverse transcribed mRNA into the genome (Vanin,

1985). Among about 20 different mammalian ribosomal protein genes charac-terized so far, the rat *S5* is the only one that seems to have no pseudogenes (Kuwano *et al.*, 1992).

In general, the vertebrate rp genes share a similar architecture. The rp genes are relatively short, the described mammalian rp genes range in length from 1.5 to 6 kb. The exon-intron structure of individual characterized rp genes is con-served within vertebrates. The only exception is *S6*, since the mouse *S6* lacks the last fifth intron present in the human homolog (Pata and Metspalu, 1996). Typical features of the rp gene promoter are the location of transcription start site (TSP) within the oligopyrimidine tract flanked by CpG-rich regions, the presence of A/T rich region instead of the canonical TATA box, short 5' leader sequence and the location of transcription regulating elements in a region of ~200 bp surrounding the TSP. A database of mammalian ribosomal protein genes with the direct link to nucleotide data and Medline references can be ac-cessed at http://www.ebc.ee/~tannilo/rp-genes.html (Annilo, 1998).

2.1.3. Regulation of expression of ribosomal protein genes

Balanced accumulation of rps is thought to be one of pre-requisites in the for-mation of functional ribosomes. In eukaryotic cells, the production of rps is co-ordinately regulated at various levels of gene expression from transcription to protein stability. In yeast, the rate at which rps are produced is controlled pri-marily at the level of transcription (Mager and Planta, 1990, Planta, 1997). In vertebrates, the rp genes are generally viewed as "housekeeping", ubiquitously transcribed genes. The most characteristic structural feature of this type of genes is location of their TSP within the area consisting of 12–25 pyrimidines. Therefore, these genes are classified as the 5' TOP (5' TerminalOligoPyri-midine) gene family. Beside the genes for rps, the 5' TOP family contains other genes involved in the production and function of the translation apparatus (re-viewed in Amaldi and Pierandrei-Amaldi, 1997).

It has been estimated that the steady state level of the single rp mRNA is about 0.1% of total cellular mRNA in normally growing cells, which constitutes up to 8% for all 80 rps (Kato *et al.*, 1994). The expression of rp mRNAs at high and comparable levels is guaranteed by the similar promoter structure of rp genes (see also section 2.1.2). Despite the similarities in overall architecture, the individual elements of distinct vertebrate rp gene promoters can be different. Among those, the binding sites for Sp1, α (RFX-1), β (GABP), γ and δ (YY-1/NF-E1/UCRBP) can be found in the promoter region of several rp genes (Hariharan *et al.*, 1989, Sáfrány and Perry, 1995). The presence of Sp1-binding GC-box sequences is common to many housekeeping and tissue specific gene promoters. Recent studies have shown that Sp1 is the one of many transcription factors acting through GC-box (Philipsen and Suske, 1999). Factor γ binds to the A/T rich region of *L32* as well as of *L30* and probably takes part in tran-

scription initiation similarly to TATA-binding protein (Yoganathan *et al.*, 1993). The factors α, β and δ are involved in transcription regulation of rp genes and several other genes (Sáfrány and Perry, 1995). It seems that the combinations of various elements create promoters with comparable efficiency and allow production of distinct rp mRNAs in similar amounts.

Although the rp genes are considered to be ubiquitously expressed genes, the level of rp mRNAs varies in different tissues, through the development (Mazuruk *et al.*, 1996) and during cell differentiation (Agrawal and Bowman, 1987). However, transcriptional control seems not to be sufficient in rapidly changing physiological conditions. Available data strongly suggest that the main regulatory mechanism of vertebrate rp genes and other 5' TOP genes is the translational control of mRNA utilization (Amaldi and Pierandrei-Amaldi, 1997). 5' TOP genes are translationally regulated in response to various physiological stimuli and inhibitors (Meyuhas *et al.*, 1987, Kaspar *et al.*, 1990, Avni *et al.*, 1994), in differentiation (Agrawal an Bowman, 1987) and in early development (Amaldi *et al.*, 1989). In all described systems, translational control of 5' TOP mRNAs involves the rapid utilization of the mRNAs in polysomes or in translationally inactive ribonucleotide particles (mRNPs) according to protein synthesis activity of cells. In actively growing cells 5' TOP mRNAs are mostly localized in polysomes whereas in quiescent or growth-inhibited cells 80% of 5' TOP mRNAs are present in mRNPs (Meyuhas *et al.*, 1987, Kaspar *et al.*, 1990). Translational regulation is mediated primarily by the 5' TOP sequence of mRNAs (Levy *et al.*, 1991). In addition to 5' TOP element, the adjacent downstream 5' untranslated region (5' UTR) is functional in translational regulation (Avni *et al.*, 1994). Depending on the sequence of 5' UTR, the translational regulation of some 5' TOP mRNAs is not ubiquitous but tissue specific. For instance, the elongation factor 2 mRNA is translationally regulated in a growth-dependent manner only in cells of hematopoietic origin (Avni *et al.*, 1997). Several proteins, including p56 in mouse, La, cellular nucleic acid binding protein CNBP and Ro60 in *Xenopus* have been shown to interact specifically with 5' UTR of 5' TOP mRNAs (reviewed in Amaldi and Pierandrei-Amaldi, 1997, Pellizzoni *et al.*, 1998). The mechanisms how these proteins are involved in translational regulation remain to be elucidated. Recently, the activation of 70 kDa S6 kinase (p70^{S6K}) and consequent phosphorylation of S6 have been suggested to play a central role in translational regulation of 5' TOP mRNAs (Jefferies *et al.*, 1994, Jefferies *et al.*, 1997).

2.1.4. Functions of ribosomal proteins

It is generally agreed that rRNAs are responsible for the basic biochemistry of protein synthesis (Green and Noller, 1997). During several years the ribosomal proteins have been viewed only as a later evolutionary embellishment which main functions are to facilitate the folding of rRNA and to maintain an optimal ribosome configuration. However, during the last decade it has been revealed

13

that rps may be active in the regulation of translation as well as they could have other functions apart from protein synthesis (reviewed in Wool *et al.*, 1995, Wool, 1996).

As mentioned in section 2.1.3., the rp S6 plays probably a role in translational regulation. Inactivation of S6 kinase p70^{S6K} with rapamycin results in decreased translation and leads to G1 cell-cycle arrest in many cell types (reviewed in Chou and Blenis, 1995). Moreover, it has been demonstrated that rapamycin selectively suppresses mitogen-induced 5' TOP mRNA translation via inhibition of p70^{S6K} (Jefferies *et al.*, 1994, Jefferies *et al.*, 1997). The only known substrate for p70^{S6K} is S6. Therefore, it has been suggested that the phosphorylation state of S6 has a central role in regulation of the translation of 5' TOP mRNAs, probably by altering the affinity of the ribosome for this type of mRNAs (Jefferies *et al.*, 1997, Amaldi and Pierandrei-Amaldi, 1997).

The presence of rat rp L10 homolog Qsr1p in yeast is required for joining 40S and 60S subunits that suggests that Qsr1p is essential for translation and therefore could take part in translational regulation (Eisinger et al., 1997). The rp L7 is involved in translational control of distinct mRNAs, most probably in extraribosomal way. L7 carries a bZIP sequence motif that mediates the formation of L7 homodimers which inhibit the translation of mRNAs through interacting with their specific sites (Neumann *et al.*, 1995). In addition, the constitutive expression of extragenic L7 in T-lymphoma cells leads to cell cycle arrest and to the apoptotic death of most of the cells, which propose that constitutive expression of L7 could specifically inhibit translation of anti-apoptotic proteins (Neumann and Krawinkel, 1997). The rp L18 has an ability to prevent double-stranded RNA-activated protein kinase dependent phosphorylation of eIF-2α. As the phosphorylation of eIF-2α has a pivotal role in growth inhibition, the overexpression of L18 has been suggested to promote protein synthesis and cell growth in certain cancerous tissues (Kumar *et al.*, 1999).

The noteworthy examples of the rps that can modulate cellular processes not associated with translation, are rps S3 and P0. Both these proteins possess apurinic/apyrimidinic endonuclease activity and have been suggested to participate in DNA repair in *Drosophila melanogaster* (Yacoub *et al.*, 1996 a, b). In addition, the rp P0 is known to be involved in the regulation of gene expression in a yet unknown manner, indicating that P0 could be a multifunctional protein (Frolov and Birchler, 1998).

Differential overexpression of several rp genes has been observed in various cancers. For example, the increased levels of *S2* mRNA have been reported in the *ras*-transformed human teratocarcinomas as well as in head and neck cancers (Chiao, *et al.*, 1992). A number of rp genes (*L31, S3, S6, S8, S12, L5* and *P0*) are overexpressed in adenomatous polyps and/or colorectal cancer (Chester, *et al.*, 1989, Pogue-Geile *et al.* 1991). Similarly, the overproduction of several rp gene (*L7a, L23a, L30, S14, S18* and *L37*) mRNAs has been observed in prostate-cancer cell lines and/or tissue samples (Vaarala *et al.*, 1998). It can be presumed that in rapidly growing and dividing cells the genes encoding proteins

14

for translational machinery are transcriptionally up-regulated. However, since the expression of vertebrate rp genes is controlled mainly at the level of translation (Amaldi and Pierandrei-Amaldi, 1997), the amounts of rp mRNAs do not reflect quantities of respective proteins directly. If not to take into consideration the extraribosomal functions of rps, it could be expected that differential overproduction of rp mRNAs in tumor cells does not contribute transformation specifically and higher (although variable) levels of mRNAs of translational machinery components could be viewed as pre-requisite for rapid protein synthesis and cell growth.

In haploid *Saccharomyces cerevisiae*, disruption of both copies of most tested rp genes one by one results in lethal phenotype. In few cases, yeast cells can survive the loss of individual rp, although with substantially reduced growth rate. Similarly, the yeast spores that carry disruption in one of two gene copies of single rps often grow at slower rate than wild type spores (reviewed in Warner and Nierras, 1998).

In *Drosophila melanogaster*, the mutations affecting rp genes often lead to *Minute* phenotype. The *Minute* mutations exhibit distinctive dominant complex phenotypes including delayed larval development, decreased fertility and viability, reduced body size and thin bristles. It has been suggested, that *Minute* phenotype occurs due to haploinsufficiency of a single ribosomal protein gene that limits production of ribosomes (Sæbøe-Larssen *et al.*, 1997, Warner and Nierras, 1998). Over 50 *Minute* loci are scattered thoroughout the *Drosophila* genome, mutations in ten rp genes (*rp 49, S2, S3, S5, S6, S13, L9, L14, L19* and *S3a*) have been shown to result in *Minute* so far (Sæbøe-Larssen *et al.*, 1997 and references therein, Beest *et al.*, 1998). However, a possibility remains that mutations in genes other than rp genes may lead to the *Minute* phenotype. At the same time, not every mutation in any rp gene results in *Minute*.

While mutations in individual rps lead to similar phenotypes in *Drosophila* as well as in yeast, there is little known about influence of mutated rp genes in vertebrates and mammalians. Considering the complexity of higher eukaryotic organisms, it has been predicted that genetic defects in ribosomal components would invariably result in early embryonic death. However, the potential role of rps in human disorders has received considerable attention recently. Human rp S4 has been shown to have two active gene copies that are located on the chromosomes X and Y. The products of *S4* genes, S4X and S4Y differ at 19 positions of the total of 263 aa and function interchangeably in ribosomes (Fisher *et al.*, 1990). It has been suggested, that haploinsufficiency of S4 contributes to the Turner syndrome (Fisher *et al.*, 1990, Kenmochi *et al.*, 1998). More recently, the rp *S19* gene have been found to be mutated in a subset of Diamond-Blackfan anaemia patients (Draptchinskaia *et al.*, 1999). The authors hypothesize that the phenotype associated with Diamond-Blackfan anaemia may result from insufficient functioning of mutated S19 in the ribosome and/or altered extra-ribosomal functions of mutated S19 in erythropoiesis and embryogenesis.

15

2.1.5. Ribosomal protein S3a and its function(s)

Mammalian S3a was first described by Collatz *et al.* (1977) as one of three ribosomal proteins (S3, S3a and S3b) that migrate to the same spot on a two-dimensional gel. After sequencing of the rat 40S subunit ribosomal proteins, it was revealed that ribosomal protein S3b has the same aa sequence as rat S3a except for the last 12 aa missing at the carboxy-terminus. In the absence of other evidences, the authors suggest that S3b results from proteolytic processing of S3a (Chan *et al.*, 1996). In the ribosome, the S3a protein is a component of the 40S ribosomal subunit domain required for initiation of translation. S3a specifically binds to initiation factors eIF-2 and eIF-3, as well as initiator tRNA, 18S rRNA and mRNA (Westermann and Nygard, 1983; Lutsch *et al.*, 1990; Bulygin *et al.*, 1997). There are no data about location of S3b in 40S subunit.

In *Saccharomyces cerevisiae*, the *S3a* homologous genes have been cloned and sequenced initially as *MFT1* and *MFT2* (Garret *et al.*, 1991), and as *PLC1* and *PLC2* genes (Kodama *et al.*, 1991). First, it was reported that the *MFT1* gene complements temperature-sensitive phenotype and is involved in targeting proteins to the mitochondria (Garret *et al.*, 1991). By contrast, Ito *et al.* (1992) showed that yeast spores with disruption of either *PLC1* or *PLC2* grow at slower rate than wild-type spores, but they do not exhibit temperature-sensitive phenotype. Afterwards, it has been demonstrated that *MFT1* locus contains two genes, *YML062* that actually is involved in mitochondrial protein import and *PLC2* gene that is homologous to mammalian S3a (Ito and Komamine, 1993, Beilharz *et al.*, 1997). According to new nomenclature, the yeast homologs of rp S3a are designated as S1A and S1B (Mager *et al.*, 1998).

In *Drosophila melanogaster,* antisense suppression of *S3a* gene disrupts ovarian development (Reynaud *et al.*, 1997) and insertion of Doc retroposon in the *S3a* gene promoter results in *Minute* phenotype (Beest *et al.*, 1998). Thus, the deficiency of S3a protein influences growing and development of *Drosophila* and yeast likewise it is described for a number of other rps. This accords with the general paradigm, that lack of single ribosomal protein prevents the biogenesis of the ribosome, and therefore the deficiency of different ribosomal proteins has a common effect in cell growth. However, since the list of rps that have functions other than structural in the ribosome has been lengthened remarkably during recent years (see section 2.1.2.), possible regulatory and/or extraribosomal function(s) of S3a have also been under discussion in several studies. The *S3a* homolog in plants (*cyc07*) was isolated from synchronously cultured *Catharanthus roseus* cells as a gene that is expressed preferentially and transiently at the G_1/S boundary (Kodama *et al.*, 1991). The promoter of *cyc07* has been shown to direct specifically expression of a β-glycuronidase reporter gene in meristemic tissues in transgenic plants (Ito *et al.*, 1994). Correspond-

ingly, it has been suggested that *cyc07* has a crucial role in cell cycle progression, cell proliferation (Kodama *et al.*, 1991, Ito *et al.*, 1994), and could take part in regulation of cell cycle and DNA replication in extraribosomal way (Chan *et al.*, 1996). As in plant cells, the *S3a* mRNA accumulates preferentially in the S-phase in synchronously cultured human fibroblasts (Kho *et al.*, 1996). However, there is no evidence that S3a functions as a regulator of cell cycle progression.

Enhanced *S3a* mRNA levels are observed in rapidly growing and dividing cells and in cells induced to proliferate. For instance, the mouse *S3a* cDNA (designated as TU-11) has been originally isolated from a library enriched for genes induced by tumor necrosis factor (Gordon *et al.*, 1992). Elevated expression of *S3a* at mRNA level has been described in several rat and human tumor cell lines and mouse hepatoma cells (Kho *et al.*, 1996, Lecomte, *et al.*, 1997). The *S3a* gene together with rp genes *L41*, *S4* and *S17* is also overexpressed in malignant lymphomas of domestic cat (Starkey and Levy, 1995). It is not known, whether the overproduction of rp S3a mRNA in these tumours contributes to the transformation of the cells.

A few studies suggest that higher expression of *S3a* gene supports the maintenance of the transformed phenotype, or in certain conditions, it could even induce the cell transformation. The rat *S3a* gene has been initially isolated as v-*fos* transformation effector gene (*fte-1*). In v-*fos* transformed Rat-1 fibroblast cells, the monoallelic disruption of *fte-1* gene with a marker plasmid led to the loss of the transformed phenotype. Exogenous overexpression of the *fte-1* gene restored the transformed phenotype of revertant cells (Kho and Zarbl, 1992). Later, the disruption of *fte-1* gene has been shown to result in decreased rate of protein synthesis and altered polysome profile, but ectopic overexpression of *fte-1* in normal Rat-1 fibroblasts failed to induce cell transformation. Two possible mechanisms whereby the reduced level of rp S3a interferes v-*fos* transformation have been suggested. First, the deficiency of S3a could prevent normal ribosome biogenesis, the decreased level of functional ribosomes reduces overall rate of protein synthesis that subsequently inhibits cell growth and causes loss of transformed phenotype. Second, the decreased S3a expression could selectively reduce synthesis of the proteins important for v-*fos* transformation compared to the overall protein synthesis rate (Kho *et al.*, 1996).

The human *S3a* gene, named as *nbl* gene has been isolated due to its abundance in Namalwa Burkitt Lymphoma cDNA library by Naora *et al.* (1995). Initially, the *nbl* expression has been noticed to be higher in tissues that exhibit internucleosomal DNA cleavage characteristic to apoptosis (Naora *et al.* 1995). Later, the suppression of *S3a* expression has been shown to induce apoptosis in cell lines that constitutively express *S3a* mRNA at high level, but not in cells expressing *S3a* mRNA at low level. In addition, the exogenous enhancement of *S3a* mRNA expression has been found to induce transformation in NIH 3T3 cells, but only if the cells have been cultured at high density. The authors con-

clude that S3a has a role in apoptotic process and in certain conditions and/or in presence of specific signals S3a could be involved in induction of transformation (Naora *et al.* 1998). It has been speculated that S3a functions in apoptosis as a translational modulator through specific inhibition of survival/protective modulators or influences apoptotic process in an extraribosomal way (Naora *et al.* 1999).

In addition to the studies outlined above, cDNAs corresponding to S3a protein gene have been cloned and characterized in *Candida albicans* (initially named as *RP10* gene) (Swoboda *et al.*, 1995), in *Oryza sativa* (Kidou *et al.*, 1994), in *Aplysia californica* (Auclair *et al.*, 1994) and in *Eimeria tenella* (Ouarzane *et al.*, 1998).

2.2. Small nucleolar RNAs

2.2.1. Post-transcriptional processing and modification of rRNA

In the nucleolus, more than 100 tandemly repeated units of ribosomal DNA (rDNA) are transcribed by RNA polymerase I into primary precursor transcripts (as 47S pre-rRNAs in mammalians). Each pre-rRNA molecule contains sequences for 18S, 5.8S and 28S rRNAs, plus additional internal (ITS1, ITS2) and external (5' ETS, 3' ETS) transcribed spacer sequences (Fig. 1). In general, mature rRNAs are produced by a series of coordinated endonucleolytic cleavages followed by exonucleolytic trimming. However, the order and intermediates generated during rRNA processing may vary among eukaryotic species as well as during growth and development in the same species (reviewed in Eichler and Craig, 1994).

Maturating pre-rRNA also undergoes extensive site specific modifications, mostly 2'-O-ribose methylation of riboses and conversion of uridine residues to pseudouridine (Fig. 2A and 2B) (Maden, 1990). The ribose methylations and pseudouridylations occur very rapidly on the pre-rRNA during or immediately after transcription. Most of the modification sites are located in the universally conserved cores of mature rRNA and are conserved among the vertebrates. A number of modifications are also conserved between vertebrates and yeast, although there are ~210 modified nucleotides identified in vertebrates and ~110 in yeast rRNA (Maden, 1990). The rRNAs of prokaryotes contain significantly less modified nucleotides. Only four 2'-O-ribose methylated nucleotides and ~10 pseudouridines have been found in *Escherichia coli* rRNA (Smith *et al.*, 1992, Ofengand and Bakin, 1997).
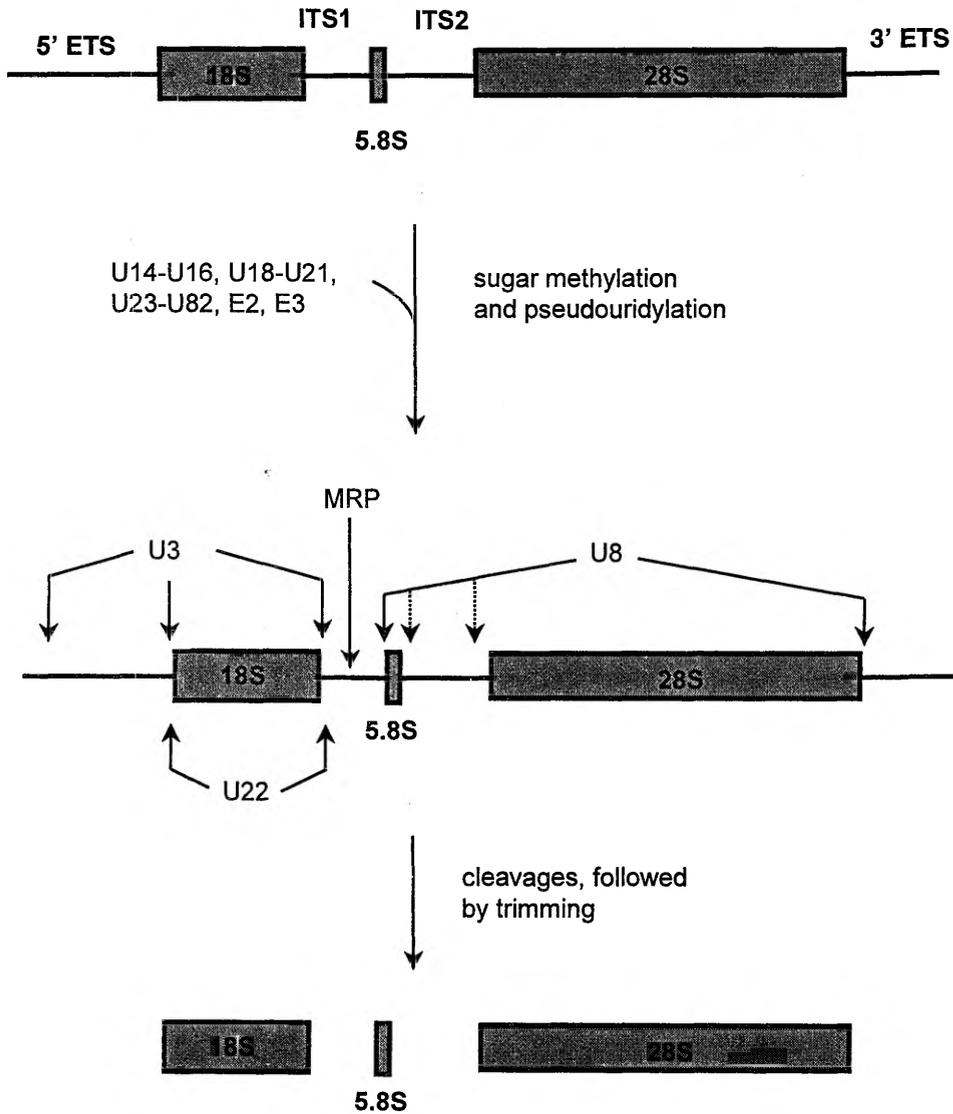
Figure 1. Modification and processing of mammalian pre-rRNA (according to Yu *et al.*, 1999). A large number of snoRNPs participate in ribose methylation and pseudouridylation of the nascent pre-rRNA transcript. Following modification, other snoRNPs and individual proteins perform the cleavages to generate immediate precursors that are trimmed to mature rRNAs. Internal and external transcribed spacers and involvement of snoRNPs are indicated. Participation of some individual proteins and potential role of vertebrate (U14, U17, E2, E3 and U13) and yeast (U14 and snR30) snoRNPs are not shown, but are discussed in the text.

19
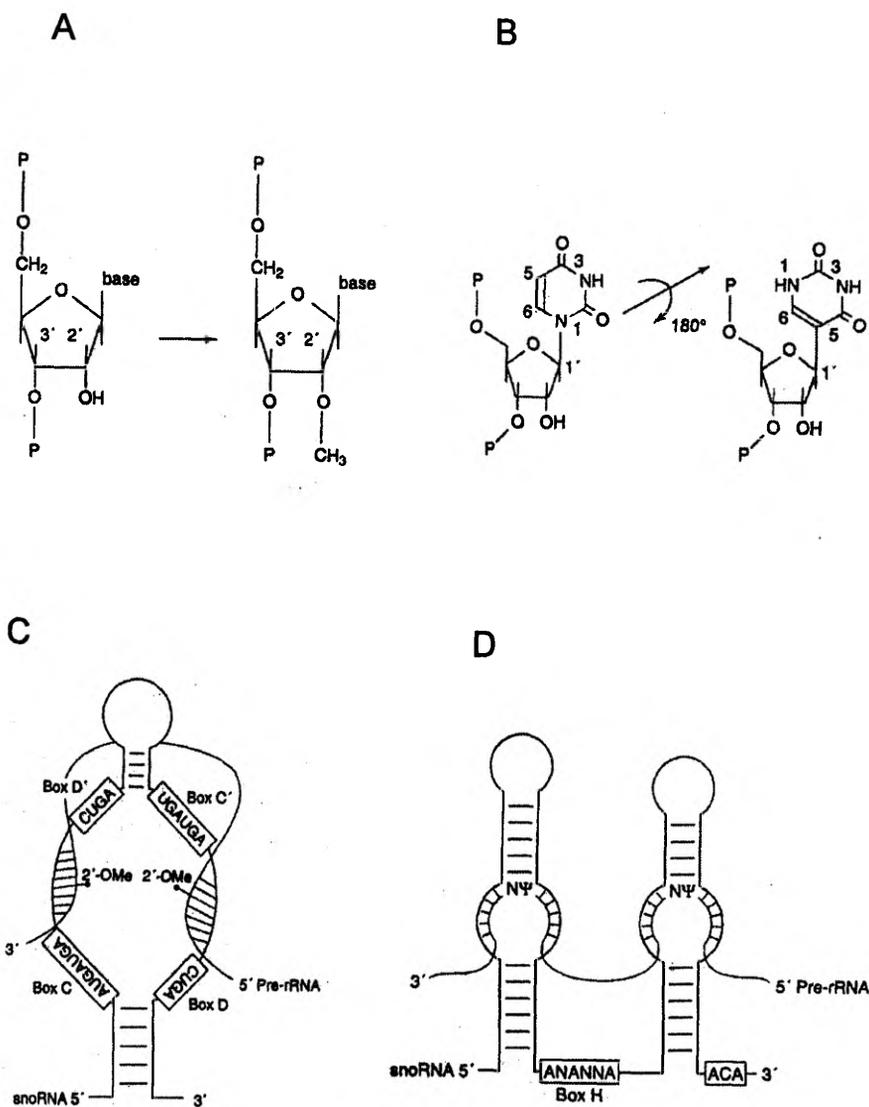
Figure 2. 2'-O-methylation, pseudouridine formation and predicted structures of the hybrids between pre-rRNA and modification guide snoRNAs (adapted from Lafontaine and Tollervey, 1998). **(A)** 2'-O-methylation of RNA nucleotides. **(B)** ψ formation by base rotation around N_3-C_6 axis. **(C)** The predicted structure of the hybrid between C/D box snoRNAs and pre-rRNA. Each box C/D and/or C'/D'-pre-rRNA interaction generates a conserved structure with the D or D' element placed five base pairs from a site of 2'-O-methylation. The base-pairing extends 10–21 consecutive nucleotides across the site of modification. Box D is implicated as the binding site for a snoRNP component, so this interaction places a common protein at a fixed position in respect to the methylation site. This positional information is presumably used by the methylase to select the 2'-O-hydroxyl group. **(D)** The predicted structure of the hybrid between H/ACA snoRNA and pre-rRNA. The sequences flanking the site of modification are base-paired to the snoRNA in a complex structure. The base-pairing involves two short stretches of 3–10 nucleotides. The base that has to be modified remains free to interact with the ψ synthase.

20

The proposed functions of ribosomal pseudouridines and 2'-O-ribose methylations remain largely speculative. In general, these modifications seem to be nonessential for viability of organisms. The unmethylated ribosomes are still formed in mammalian cells treated with cycloleucine, an inhibitor of S-adenosyl-methionine production. However, the overall efficiency of rRNA processing is substantially decreased and its kinetics is delayed during the cycloleucine treatment (Caboche and Bachellerie, 1977). Therefore, methylation has been proposed to have a role mainly in facilitating ribosomal subunit assembly (Bachellerie and Cavaillé, 1998). Since ribose methylations increase hydrophobicity in the positions they occur, it has been suggested that they also may stabilize rRNA (Maden, 1990). Interestingly, the ribose methylation level of archaeal hyperthermophile *Sulfolobus solfataricus* rRNA has been observed to increase in higher temperatures. In addition, an overall number of 2'-O-methyl groups in rRNA of *Sulfolobus* is remarkably higher compared with eubacterials. These data also support the hypothesis that ribose methylations are implicated in stabilization of rRNA (Noon *et al.*, 1998).

The knowledge of the role of ribosomal pseudouridines is very limited as well. If the pseudouridine forms theoretical Watson-Crick base pair with adenine, conformational differences occur compared with a standard A-U pair. Pseudouridine confers also an extra hydrogen bonding capacity compared with uridine. Therefore, it has been proposed that pseudouridines somehow contribute to ribosome assembly in some manner and/or could take part in ribosomal functions (Maden, 1990, Ganot *et al.*, 1997).

A number of nonribosomal proteins assist rRNA processing and ribosome assembly. Probably a major and most studied nucleolar protein is nucleolin. It has been demonstrated that the vertebrate nucleolin functions in the first step of rRNA processing, i. e. cleavage within 5' ETS (Ginisty *et al.*, 1998). In addition, nucleolin is implicated in rDNA transcription, ribosome assembly, nucleocytoplasmic transport and chromatin structure (reviewed in Ginisty *et al.*, 1999). Several ribonucleases and RNA helicases that are required for rRNA processing in vertebrates or yeast have been isolated. For instance, mammalian endonuclease RNase PR1 cleaves the pre-rRNA within 5' ETS, the 5'–3' exonucleases Rat1p and Xrn1p are needed for maturation of the 5.8S rRNA, the helicases DSR1 and SPB4 are necessary for normal production of the 28S rRNA (reviewed in Eichler and Craig, 1994), Rnt1p, a yeast homolog of prokaryotic RNase III, has been implicated in cleavages within 5' ETS and 3' ETS (Abou Elela *et al.*, 1996). The 3'–5' exonucleolytic processing of yeast 5.8S is performed by five proteins (Rrp4p, Rrp41p, Rrp42p, Rrp43p and Rrp44p) that from a complex called exosome. Among them, Rrp43p has been shown to be required also for maturation of 28S and 18S rRNAs (Zanchin and Goldfarb, 1999).

In addition to the individual proteins, a number of small nucleolar ribonucleotide particles (snoRNPs) are involved both in pre-rRNA processing and modification (Fig. 1). In general, the snoRNPs consist of a RNA component

(60–300 nt) called snoRNA and several proteins. During the past few years, it has been recognized that a majority of the snoRNPs (up to 200 in vertebrates) act as guide RNAs in the nucleotide modifications of pre-rRNA and only a few of them orchestrate pre-rRNA cleavage and folding (reviewed in Smith and Steitz, 1997). The structure of snoRNAs, snoRNP proteins and current understanding of the role of snoRNPs in rRNA maturation are discussed more thoroughly in the next four sections.

2.2.2. Structure and classification of snoRNAs

More than 60 snoRNA genes, although not always homologous, have been determined both in vertebrates and yeast (reviewed in Yu *et al.*, 1999, http://www.bio.umass.edu/biochem/rna-sequence/Yeast_snoRNA_Database/snoRNA _DataBase. html). Almost all known snoRNAs can be divided into two groups on the basis of the conserved sequence elements they contain. A large group of snoRNAs is defined by the sequence elements as box C (UGAUGA) and box D (CUGA) (Fig. 2C) (Maxwell and Fournier, 1995, Nicoloso *et al.*, 1996). The C and D boxes are positioned close to the 5' and 3' ends of a snoRNA, respectively, and they are frequently brought together by short complementary sequences found in close proximity to these elements. The middle part of a snoRNA is generally less structured and may contain imperfect copies of the authentic C and D boxes, referred to as C' and D' boxes (Maxwell and Fournier, 1995, Kiss-László *et al.*, 1998). In addition, each C/D box snoRNA possesses one or two antisense element(s) of 10–21 nt of complementarity to mature rRNA (reviewed in Bachellerie and Cavaillé, 1998). It has been demonstrated that C and D boxes are essential for binding the snoRNP proteins as well as for processing and stability of this type of snoRNAs (Caffarelli *et al.*, 1996, Watkins *et al.*, 1996, Lafontaine and Tollervey, 1999).

Another major class of snoRNAs has an evolutionarily conserved (hairpin-hinge-hairpin-tail) secondary structure (Fig. 2D). The single-stranded hinge region and the short 3'-end tail contain two conserved sequence elements H (AN-ANNA) and ACA, respectively (Balakin *et al.*, 1996, Ganot *et al.*, 1997a). Like C and D box elements, the boxes H and ACA are required for accumulation and stability of this class of snoRNAs (Ganot *et al.*, 1997b, Bortolin *et al.*, 1999).

There is still another type of snoRNAs: RNase MRP and RNase P. The RNA components of these two snoRNPs can be folded into very similar cage-shaped secondary structures (Forster and Altman, 1990). In addition, RNase MRP and RNase P contain several identical proteins (see section 2.2.4.)(Pluk *et al.*, 1999 and references therein).

22

2.2.3. Functions of snoRNAs

As snoRNAs are located in the nucleolus, they are all expected to take part in ribosome biogenesis. A majority of snoRNAs participate in modification events of the rRNA. Accordingly, most of the antisense C/D box snoRNAs are required for site-specific 2'-O-ribose methylation of pre-rRNA (Nicoloso *et al.*, 1996, Kiss-László *et al.*, 1996, Cavaillé *et al.*, 1996a). One snoRNA can take part in methylation of one or two riboses. For instance, 41 known methylation guide snoRNAs participate in methylation of 51 rRNA riboses in *Saccharomyces cerevisiae* (Lowe and Eddy, 1999). For each site of 2'-O-methylation in the mature rRNA the antisense element of specific snoRNA pairs with the corresponding region of the pre-rRNA. Analysis of the box C/D snoRNA base-pairing regions has shown that the sites of rRNA modification are precisely positioned with respect to the conserved D or D' box element, the fifth nt upstream D or D' box faces invariably the nucleotide that will be methylated in the rRNA (Fig. 2C). Accordingly, methyltransferase has been assumed to interact directly or indirectly with D or D' boxes, and to use this positional information to select the target nt for the methyltransfer reaction (Nicoloso *et al.*, 1996, Kiss-László *et al.*, 1996). In addition, the presence and correct spacing of C or C' box motifs have been demonstrated to be required for efficient rRNA methylation (Kiss-László *et al.*, 1998). It is not known, how and by which enzyme(s) the methyltransfer reactions are carried out in eukaryotic rRNA (see also section 2.2.4.).

In yeast, two snoRNAs have been shown to guide two adjacent methylations. Disruption of U24 snoRNA results in loss of the 2'-O-methyl group at predicted position Am1447, as well as the neighbouring Am1448 (Kiss-László *et al.*, 1996). Similarly, disruption of snR13 leads to loss of 2'-O-methylation at both, the predicted Am2278 and adjacent Am 2279 (Lowe and Eddy, 1999). It has been hypothesized that an independent methyltransferase could catalyze the 3' adjacent methyl groups (Kiss-László *et al.*, 1996). Alternative proposal suggests that a change in the snoRNA-rRNA base pairing could allow a single snoRNA to guide modification at the observed 3' adjacent 2'-O-methyl site (Lowe and Eddy, 1999).

Other snoRNAs that possess C and D boxes (U3, U8, U22, and U13) are not expected to direct 2'-O-methylation. U3 snoRNA has been demonstrated to pair functionally with multiple sites of pre-rRNA within 5' ETS, 18S rRNA and within ITS1 (Beltrame and Tollervey, 1995, Hughes, 1996). Accordingly, U3 snoRNA has been shown to be essential for pre-rRNA cleavages within 5' ETS and ITS1, and is required for synthesis and, probably, folding of 18S rRNA, both in vertebrates and yeast (Borovjagin and Gerbi, 1999, Hughes and Ares, 1991, Sharma and Tollervey, 1999). U8 snoRNA has been shown to be essential for cleavage at the 3'-end of 28S rRNA and upstream of 5.8S rRNA in *Xenopus* (Peculis and Steitz, 1993). At the same time, U8 has complementarity to 5.8S-ITS2 junction and 5'-end of 28S RNA and could organize a functional secondary structure pocket containing cleavage sites for ITS2 excision (Peculis, 1997,

Michot *et al.*, 1999). Depletion of U22 from *Xenopus* oocytes has been shown to prevent the processing of 18S rRNA at both ends (Tycowski *et al.*, 1994). U13 contains a complementarity to 18S rRNA within 3' terminal segment that is essential for rRNA processing and, therefore, U13 could act in processing of 18S/ITS1 junction (Cavaillé *et al.*, 1996b).

U14 differs from all other C/D box snoRNAs since it possesses two regions, termed domains A and B, with complementarity to two sequence elements that are located in close proximity in the secondary structure of 18S rRNA (Tollervey and Kiss, 1997). In *Xenopus*, it has been shown that domain B functions as a guide in methylation of 18S rRNA (at nt 427), whereas domain A has been found to be nonessential for 18S rRNA maturation (Dunbar and Baserga, 1998). At the same time, U14 has been shown to facilitate the first processing event in mouse rRNA maturation, cleavage within the 5' ETS *in vitro* (Enright, 1996). In yeast, domain A of U14 appears to be required for cleavages of pre-rRNA within the 5' ETS and ITS1 (Liang and Fournier, 1995). In addition, U14 of *Saccharomyces cerevisiae* contains a domain Y that is conserved in yeast, but is absent in vertebrates. It has been reported that domain Y cross-links to pre-rRNA at several sites and suggested that it can play a role in 18S rRNA processing (Samarsky *et al.*, 1996, Morrisey and Tollervey, 1997).

Majority of H/ACA snoRNAs direct the site-specific synthesis of pseudouridine (ψ) in pre-rRNA (Ganot *et al.*, 1997a, Ni *et al.*, 1997). While methylation guide snoRNAs possess long uninterrupted complementary sequences to rRNA, the ψ guide snoRNAs contain two shorter complementary motifs. These two recognition motifs are brought together by a stem-loop structure and they pair with rRNA sequences preceding and following a pseudouridylation site. Between these sequences, two rRNA nucleotides remain unpaired, the first one of them (in a 5' to 3' orientation in rRNA) is invariably a uridine that is converted to ψ (Fig. 2D) (Ni *et al.*, 1997; Ganot *et al.*, 1997b). Pseudouridylation reaction is carried out by pseudouridine synthase (Cbf5p in *Saccharomyces cerevisiae*), that is one of four common known H/ACA snoRNP proteins (Lafointaine *et al.*, 1998, Watkins, *et al.*, 1998a). For some H/ACA snoRNAs (snR5, snR8, snR44) it has been experimentally demonstrated that they contain two distinct pseudouridylation pockets (Ni *et al.*, 1997; Ganot *et al.*, 1997b). It has been suggested, that all H/ACA snoRNAs could have two active ψ synthesis centres, but the second substrate for most snoRNAs has not been identified yet (Watkins, *et al.*, 1998a). An interesting hypothesis for ψ guide snoRNAs function has been suggested by Ofengand and Fournier (1998). They suggest that pseudouridylation snoRNAs could actually function as chaperones in the pre-rRNA folding process, and ψ formation would be merely a benign byproduct designed as a signal to tell the guide RNA to dissociate from the pre-rRNA once its folding job is done.

So far, only two yeast H/ACA snoRNAs, snR10 and snR30, have been found to be essential for normal pre-rRNA processing and cell growth in yeast (Tol-

lervey, 1987, Morrissey and Tollervey, 1993). Genetic depletion of other individual H/ACA snoRNAs results in lack of respective ribosomal pseudouridines, but no phenotype different from wild type cells has been detected (reviewed in Tollervey and Kiss, 1997, Ganot et al., 1997a). However, a few studies indicate that H/ACA snoRNAs may take part in cleavage and folding of rRNA as well. Two H/ACA snoRNAs U17(E1) and E3 have been demonstrated to facilitate first cleavage within the 5' ETS in mouse cell extract in vitro (Enright et al., 1996). Oligodeoxynucleotide targeting experiments in Xenopus oocytes have shown that U17(E1), E2 and E3 may function in rRNA processing at following cleavage sites: E1 at the 5' end and E2 at the 3' end of 18S rRNA, and E3 near the 5' end of 5.8S rRNA (Mishra and Eliceiri, 1997). On the other hand, E2 and E3 (but not U17) have been proposed to function as guide snoRNAs in pseudouridylation of the human 28S rRNA (Ganot et al., 1997a). In yeast, H/ACA snoRNA snR30 seems to be similar to vertebrate U17 in that it does not appear to be involved in pseudouridylation. The cross-linking and depletion experiments have shown that snR30 interacts with the pre-rRNA and could participate in 18S rRNA maturation (Morrissey and Tollervey, 1993).

Thus, the role of non-modification C/D and H/ACA snoRNAs in ribosome biogenesis is not always completely understood. In general, these snoRNAs seem to be required for establishing the correct structure for pre-rRNA processing and ribosome assembly, rather than they directly participate in cleavage reactions themselves. Accordingly, yeast endoribonuclease Rnt1p (homologous to bacterial RNaseIII) has been shown to cleave pre-rRNA within 3' ETS and also at a U3 snoRNA-dependent site within 5' ETS that excludes catalytic role of snoRNAs in these sites (Abou Elela et al., 1996). It is also possible that some snoRNAs could act as modification guides of other RNAs with nucleolar localization. Recently, the depletion of mgU6-77 snoRNA from Xenopus oocytes has been shown to prevent 2'-O-methylation of both the C77 nt in U6 and of the C2970 nt in 28S rRNA, methylation can be restored by injecting in vitro transcribed mgU6-77 (Tycowski et al., 1998).

Two remaining snoRNPs, RNase MRP and RNase P are exceptional, since they can function as independent endoribonucleases. RNase MRP has been originally identified as an endoribonuclease that cleaves an RNA substrate derived from the mitochondrial origin of DNA replication in vitro (Chang and Clayton, 1987a, 1987b). However, the vast majority of RNase MRP is localized in the nucleolus, where it is involved in the processing of the 5' end of 5.8S rRNA. RNase MRP has been demonstrated to cleave the yeast pre-rRNA within ITS1 in vivo and in vitro (Schmitt and Clayton, 1993, Chu et al., 1994, Lygerou et al., 1996). The RNA component and all characterized RNase MRP proteins are required for cell growth. However, the primary function of RNase MRP is still unclear, as neither cleavage within ITS1 nor preparation the mitochondrial replication primers is essential for viability (Tollervey and Kiss, 1997).

RNase P is known mainly as an endoribonuclease responsible for 5' maturation of precursor tRNAs in all organisms (reviewed in Frank and Pace, 1998).

RNase P has been localized both in the nucleolus and nucleoplasm in mammalian and yeast cells (Jacobson *et al.*, 1997, Bertrand *et al.*, 1998). Concordantly, the precursor tRNAs and their processing events have been localized to the nucleolus as well as nucleoplasm (Bertrand *et al.*, 1998). In addition, the RNase P RNA subunit mutations have been reported to affect rRNA processing within ITS1 and ITS2 regions in yeast (Chamberlain *et al.*, 1996).

2.2.4. snoRNP proteins

The members of each class of snoRNAs are associated with common protein components in snoRNP particles. A large number of C/D box snoRNAs have been found to bind fibrillarin (Nop1p in yeast) that most probably is a common protein to all snoRNAs of this class (reviewed in Maxwell and Fournier, 1995, Bachellerie and Cavaillé, 1998). Genetic depletion of Nop1p inhibits the cleavages of the pre-rRNA within ITS1 and ITS2 that is consistent with Nop1p association with U3 and U14 snoRNAs (Tollervey *et al.*, 1991). Certain different mutations in Nop1p have been shown to result in distinct phenotypes, exhibiting defects in pre-rRNA processing, methylation or assembly of ribosomal subunits (Tollervey *et al.*, 1993). Recently, a second common protein component (Nop5p or Nop58p) of C/D box snoRNPs has been identified and characterized in *Saccharomyces cerevisiae*. Genetic depletion of Nop58p leads to loss of all tested C/D box snoRNAs, strong inhibition of pre-rRNA processing and accumulation of 3' extended forms of U3 and U24, showing that this protein is involved in snoRNA synthesis as well (Wu *et al.*, 1998, Lafontaine and Tollervey, 1999). Human U3 snoRNP has been reported to contain at least six proteins with molecular masses of 74, 59, 36 (fibrillarin), 30, 13 and 12.5 kDa (Parker and Steitz, 1987). In addition to fibrillarin that is common to all C/D box snoRNAs, four proteins specific to U3 in yeast (SOF1, Mpp10p, Imp3p and Imp4p) and one in human (hU3-55k) have been characterized (Jansen *et al.*, 1993, Dunbar *et al.*, 1997, Lee and Baserga, 1999, Pluk *et al.*, 1998). The mouse U14 snoRNP complex has been shown to contain at least three C/D box core motif-binding proteins (50, 55 and 65-kD) and one potential U14 specific 38-kDa protein (Watkins *et al.*, 1998b). In *Xenopus*, a RNA-binding 68-kDa protein (p68) is known as a common C/D box snoRNPs component in addition to fibrillarin (Caffarelli *et al.*, 1998).

There is no indication, whether any of the identified C/D box snoRNP proteins could function as a 2'-O-ribose methylase. Moreover, it is not known whether the methylase is a snoRNP-associated protein or is it a diffusible factor that recognizes and transiently binds the pre-formed pre-rRNA-snoRNP complex. Before the participation of snoRNAs in rRNA modifications was established, a nucleolar 50 kDa protein containing 2'-O-ribose methyltransferase activity had been identified and partially purified from mammalian cells (Segal and Eichler, 1991). Future characterization of this protein would help to answer

a lot of questions about the structure and functioning of the rRNA methylation apparatus.

For H/ACA snoRNAs, four common proteins (Gar1p, Cbf5p, Nhp2p and Nop10p) have been characterized in yeast (Girard *et al.*, 1992, Henras, *et al.*, 1998, Lafontaine *et al.*, 1998). Cbf5p has been demonstrated to be involved in ribosome biogenesis and it shows striking homology with known prokaryotic pseudouridine synthases. Since Cbf5p is present in all tested H/ACA snoRNP complexes, it has been suggested that it is a common pseudouridylase in all H/ACA snoRNPs (Lafontaine *et al.*, 1998, Watkins, *et al.*, 1998a).

RNase MRP and RNase P of *Saccharomyces cerevisiae* have been shown to contain eight common proteins: Pop1p, Pop3p, Pop4p, Pop5p, Pop6p, Pop7p/Rpp2p, Pop8p and Rpp1p (Chamberlain *et al.*, 1998 and references therein). In addition, one RNase MRP specific protein (Snm1p) has been identified in yeast (Scmitt and Clayton, 1994). In HeLa cells, the RNA molecule of RNase P is associated with nine proteins: hPop1p, hPop4p, Rpp14, Rpp20, Rpp25, Rpp29, Rpp30, Rpp38 and Rpp40 (Jarrous *et al.*, 1999 and references therein, van Eenennaam *et al.*, 1999). According to current knowledge, four of them (hPop1p, Rpp30, Rpp38 and hPop4p) are also present in RNase MRP (Pluk *et al.*, 1999, van Eenennaam *et al.*, 1999).

2.2.5. Organization and synthesis of snoRNAs

Genetic organization of snoRNAs is interesting and variegated, different expression strategies are in use. In vertebrates, a few snoRNAs are transcribed from their own promoter and terminator signals by RNA polymerase II (U3, U8 and U13) or by RNA polymerase III (RNase P and RNase MRP) (reviewed in Maxwell and Fournier, 1995, Tollervey and Kiss, 1997). A large majority of vertebrate snoRNAs are encoded within introns of protein-coding genes or processed from pre-mRNA-like primary transcripts that do not generate functional mRNAs. In vertebrates, one snoRNA is invariably present per one intron. Interestingly, all established snoRNA host genes belong to the 5'-terminal oligopyrimidine (5'TOP) gene family (Smith and Steitz, 1998, Pelczar and Filipowicz, 1998). Most of these genes encode proteins that are involved in nucleolar function, ribosome structure or protein synthesis (reviewed in Tollervey and Kiss, 1997, Smith and Steitz, 1998). Four genes (*UHG, U19H, U17HG* and *gas5)* encode single or multiple snoRNAs in their intronic sequences, but do not produce any protein (Tycowski *et al.*, 1996, Bortolin *et al.*, 1998, Smith and Steitz, 1998, Pelczar and Filipowicz, 1998). As 5' TOP family genes are ubiquitously and highly transcribed genes, it has been suggested that they have been selected for snoRNA hosts to guarantee the transcription rate high enough to ensure a sufficient amount of snoRNAs (10^4 copies per cell) to base pair with the nascent pre-rRNA molecules. Besides, it could indicate that the expression of snoRNAs and the expression of genes encoding components of translational

maschinery is coregulated at transcriptional level in a specific, but yet unknown reason. In addition, the 5' TOP sequence elements have been supposed to have a role in binding the factors that later assist in processing of snoRNAs (Smith and Steitz, 1998, Pelczar and Filipowicz, 1998).

In yeast, snoRNAs can be transcribed as independent transcription units, or as polycistronic clusters where two or more snoRNAs are expressed from upstream promoter, or within the introns of protein-coding genes (Chanfreau *et al.*, 1998, Lowe and Eddy, 1999). Like in vertebrates, single snoRNA per one intron is present in yeast. All the described plant snoRNAs are synthesized from polycistronic clusters. Unlike the yeast polycistronic snoRNAs, the plant snoRNA arrays can occur either in intronic or non-intronic locations. Similarly to single intronic snoRNAs, the clusters are synthezised as part of pre-mRNA transcripts of the host protein-coding genes (reviewed in Brown and Shaw, 1998).

In vertebrates and yeast, the presence of only one single snoRNA per intron is important since synthesis of these snoRNAs is largely splicing-dependent and involves debranching of the excised intron and following exonucleolytic trimming (Kiss and Filipowicz, 1995, Cavaillé and Bachellerie, 1996c). However, two snoRNAs U16 and U18, that are both encoded by rp *L1* gene introns in *Xenopus*, are known to be synthesized splicing-independently. Moreover, processing of these snoRNAs is mutually exclusive with pre-mRNA splicing and takes place via endonucleolytic cleavages and exonucleolytic trimming in *Xenopus* (Caffarelli *et al.*, 1996, Caffarelli *et al.*, 1997).

Synthesis of polycistronic pre-snoRNAs in non-intronic as well as intronic transcripts involves also endonucleolytic cleavages and trimming with exonucleases (Leader *et al.*, 1997). In several cases, it has been shown that common proteins take part in processing of the pre-rRNAs and snoRNA precursors. Recently, the yeast ortholog of bacterial RNAseIII, (Rnt1p) has been demonstrated to function as an endonuclease in synthesis of many snoRNAs that are encoded by monocistronic or polycistronic transcription units. Rnt1p cleaves potentially double-stranded regions closed by tetraloops with the consensus sequence AGNN (Chanfreau *et al.*, 1998). Two yeast exonucleases, Xrn1p and Rat1p have been reported to carry out the 5' exonucleolytic trimming of various snoRNAs. These exonucleases are also responsible for the degradation of several excised fragments of pre-rRNA spacer regions and generating the 5' end of 5.8S rRNA (Petfalski *et al.*, 1998). Similarily, the yeast exosome (the complex that consists of multiple 3'–5' exonucleases) that is required for processing the 3' end of 5.8S rRNA has been shown to be involved in 3' processing of polycistronic as well as intron-encoded snoRNAs (Allmang *et al.*, 1999).

3. PRESENT INVESTIGATIONS AND DISCUSSION

3.1. Ribosomal protein S3a (References I and II)

3.1.1. Isolation and primary sequence of the human S3a cDNA

In the first step, the rat rp S3a-specific antibody characterized by Lutsch *et al.* (1990) was used for screening a human cDNA expression library. Sixteen clones were isolated, among which the clone with the longest insert was sub-cloned and sequenced. This clone was with the length of 721 bp, but missed the 5' part of the cDNA. To isolate a full-length *S3a* cDNA, a new library in λZA-PII vector was screened with a PCR-generated ^{32}P-labelled partial cDNA probe. After the first screening, the colonies were analysed by PCR with cDNA-specific oligodeoxynucleotide (complementary to nt 400–420 on Fig. 1b. in reference I) and T3 or T7 primers in two different sets of reactions. In this way, the 5' part of cDNA was amplified regardless of the insert orientation. DNA from two longer clones was purified and sequenced. The clone with the longest insert was found to contain two different cDNAs ligated together via their 5' ends. The clone with the second longest insert revealed 115 nt additional sequence to previously isolated cDNA (altogether 836 bp), but it did not represent the full-length cDNA. Efforts to isolate a full-length *S3a* cDNA from other libraries did not succeed either.

Northern hybridization analysis of *S3a* mRNA was performed. Using *S3a* cDNA specific ^{32}P-labelled riboprobe, one mRNA band with the length of about 900 nt was determined (Fig. 2 in reference 2). Thus, the detected mRNA length indicated that the longest isolated *S3a* cDNA could miss only the 5' noncoding part and few nt from the coding region.

3.1.2. The human S3a protein

In order to verify the sequence of the *S3a* cDNA clone, direct N-terminal amino acid sequencing of purified S3a was performed. Except the first Ala, the determined S3a sequence (Ala-Val-Gly-Lys-Asn-Lys-Arg-Leu-Thr-Lys) matched exactly to that of deduced from the cDNA sequence. Thus, the longest described S3a cDNA revealed to miss 4 nt from the coding part, including the AUG start codon (Fig. 1a in reference I). Absence of the first Met from the S3a protein sequence most probably implies that this Met is removed during post-translational processing. Accordingly, mass spectrometric analysis has demonstrated that difference between the predicted and observed rat S3a masses is consistent with its demethionylation (Louie *et al.*, 1996). Without the first Met, the human S3a protein consists of 263 aa and its calculated molecular weight is

8   29

about 29.8 kDa. Similarly to other rps, S3a is basic (it has a pI of 10.53) and contains unusually large number of lysines (38 of 263 aa).

The human rp S3a contains two sequence regions similar to nuclear localization signals, the classical type of NLS pattern at positions 219–222 and the bipartite NLS at 151–167 (predicted with program PSORTII at http://psort.nibb.ac.jp:8800/). The presence of S3a in nucleolar 80S pre-ribosome has been demonstrated with antibodies (Todorov et al., 1983). Nevertheless, it has been demonstrated that the β-galactosidase fusions holding whole S3a or various S3a parts do not have nuclear targeting ability (Hoth, 1993). Therefore, it is possible that S3a uses NLS-independent nuclear import pathway.

It has been demonstrated in various studies that mammalian rp S3a can bind to 18S rRNA, mRNA and initiator tRNA (Westermann and Nygard, 1983; Lutsch et al., 1990; Bulygin et al., 1997). As for many other rps, the precise RNA binding strategy of S3a remains unknown. The rp S3a contains no known RNA and DNA binding motifs available in motif databases (http://www.motif.genome.ad.jp/).

3.1.3. The *S3a* gene family

The copy number of the human *S3a* gene was evaluated by Southern blot analysis. Hybridization to *Bam*HI, *Pst*I, *Kpn*I, *Sac*I, *Hind*III and *Bam*HI/*Hind*III digested HeLa DNA with ^{32}P-labelled S3a cDNA probe suggests that there are 6–8 copies of *S3a* in human genome (Fig. 3 in reference I). Later, only one of these copies has been shown to be intron-containing (Nolte et al., 1996). Southern analysis of mouse genomic DNA using the *S3a* cDNA probe indicated that there are 9–11 members for *S3a* in mouse genome. With the probes specific to mouse *S3a* 4th intron, the presence of a single *S3a* intron-containing gene was detected in mouse genome (data not shown). Thus, the human and mouse *S3a* follow the general rule found for mammalian rp genes, according to which one functional intron-containing gene occurs among multiple pseudogenes.

3.1.4. Isolation and characterization
of the human and mouse *S3a* genes

In order to isolate human and mouse *S3a* genes in the presence of multiple pseudogenes, a combination of PCR and conventional library screening strategy was used as described by Feo et al. (1990). The human and the mouse genomic λ-based libraries were both divided into 36 pools, each containing 20,000 recombinant phage particles and amplified about 10^5 fold. The pools were screened with PCR using primer pairs derived from the human *S3a* cDNA (see paragraph 2.1. in reference II). The PCR resulted in two amplification products: the larger corresponding to the intron-containing gene and the shorter corresponding to the processed pseudogenes. The pools positive for the intron-

30

containing gene were plated and screened using the human and mouse ^{32}P-la-belled *S3a* cDNA probes. One λ clone containing genomic DNA for human *S3a* and one for mouse *S3a* were isolated, subcloned and the nucleotide sequence was determined for both genes together with the flanking regions. The mouse *S3a* gene was demonstrated to be approximately 4.6 and its human counterpart about 5.1 kb in length. Both genes were found to contain six exons and five introns all with identical positions (Fig. 1 and Fig. 2. in reference II).

While the present work was in progress, Nolte et al. (1996) reported the cloning and sequencing of the human S3a-encoding gene. Comparison of the published human *S3a* sequence to that obtained in our laboratory (in 5107 bp region) showed that both sequences are identical except for twelve nucleotide positions, all located in introns. It remains unclear whether the differences between the two human *S3a* independently determined sequences represent the naturally occurring polymorphisms or sequencing errors.

3.1.5. Sequence analysis of the human and mouse *S3a* promoter regions

Promoters of different rp genes share several features and transcription factor binding sites that can be also found in the human and mouse *S3a* genes. The 5' oligopyrimidine tract or 5' TOP is the most characteristic regulatory element in vertebrate rp genes. The 5' TOP of the human and mouse *S3a* gene has identical sequence (CTTCCGCCCTTTT) and is disrupted with one G. TSP of human *S3a* has been mapped to the C nt (underlined above) within the 5' TOP (Nolte *et al.* 1996). The primer extension analysis on mouse total RNA confirmed (data not shown) that mouse TSP maps to the same nt within the mouse *S3a* oligopyrimidine tract (Fig. 2 in reference II). Thus, the first exon of the human and mouse *S3a* contains 25 nt of untranslated and 62 nt of coding sequence. Concerning vertebrate rp genes, this is unusual, since normally the first exon of rp genes is either untranslated or it encodes only the first few amino acids. Among all the mammalian rp genes characterized so far, only the mouse *S16* has similar relatively large first exon (Wagner and Perry, 1985).

Like in other housekeeping genes, the 5' ends of the human and mouse *S3a* are embedded in a CpG-rich region (G/C content 64% and 60%, respectively, and CpG/GpC ratios 0.95). The promoters of both genes lack a well-defined TATA box, instead of which an A/T rich area between nt −28 to −22 is present. Concidering putative binding sites for transcription factors, the human and mouse *S3a* both contain two sites for GABP-factor with the consensus CGGAAR in either orientation (Hariharan *et al.*, 1989) at nt positions −82 to −77 and −6 to −1, and two Sp1-recognition sites (Pugh and Tjian, 1990) at nt −3 to +3 and −42 to −37. In addition, the mouse *S3a* has the sequence TCGCGAGA at nt −101 to −93 similar to the box A element in *L7a* promoter (Colombo and Fried, 1992) and three Sp1-binding sequences at nt positions

31

+140 to +145, +164 to +169, +267 to +272 within the first intron that are not found in the human homolog (Fig. 2 in reference II).

3.1.6. The S3a protein family

The S3a homologs have been identified in 23 species including vertebrates, insects, plants, fungi and archaebacteria (http://www.expasy.ch/cgi-bin/get-sprot-entry?p49241). Thus, S3a belongs to a large group of rps that can be related to arhaebacterial rps, but which have no eubacterial homologs. The amino acid sequences of seven different S3a homologs are aligned in Fig. 3.

Among characterized mammalian S3a proteins, the human and cat homologs are 100% identical to each other (Starkey *et al.*, 1996). The mouse and rat S3a are highly conserved as well, but compared to human S3a they both reveal two differences at the C-terminal region between aa 235–241 (Fig. 3). Ser235 and Thr240 in the human protein are replaced by Gly235 and Ala240 in rat S3a; in mouse, Ser235 is replaced by Gly235 and Thr241 by Ala241. In addition, computer analysis of mouse S3a ESTs (http://www.ncbi. nlm.nih.gov/dbEST/) showed that two aa from the same region of S3a are polymorphic in laboratory mouse strains. The amino acid 236 can specify Ser (this study) or Thr (found in 68% of mouse S3a ESTs), and the codon 237 may encode Ser (most frequent), as well as Arg (20%), Pro (5%), and Leu (2.5%).

The relationship between human S3a and nonmammalian homologs is weaker. For example, the *Drosophila melanogaster* S3a shows 68,5% of identity, the *Catharanteus roseus* S3a 63.2%, the *Saccharomyces cerevisiae* S1A 57.6% and S1B 58.0%, and the archaebacteria *Methanococcus jannaschii* S3a 28.9% (in 211 aa overlap) to human rp S3a. This conservation rate of rp S3a is quite typical to many rps that possess archaebacterial homologs (Wool *et al.*, 1995).

3.1.7. Genomic localization of the human *S3a* gene

Chromosomal localization of human *S3a* was determined by fluorescence *in situ* hybridization (FISH). Using fluorescent microscopy, 35 metaphase spreads hybridized with the human genomic probe were examined. Fluorescent hybridization signals were detected on one or both homologs of chromosome 4 in 89% of metaphases (Fig. 5a in reference II). No other locus with specific fluorescent spots was detected. Parallel examination of the hybridization signals and the DAPI-banded chromosomes assigned the locus for human *S3a* to the long arm of chromosome 4, to the band q31.2-3 (Fig. 5b,c in reference II). Using intron-specific STS mapping, two other rp genes (*L34* and *L9*) have been located on chromosome 4 in addition to *S3a* (Kenmochi *et al.*, 1998). Chromosomal localization of mouse *S3a* has not been determined. According to Davis Human/Mouse Homology Map, mouse *S3a* is located on chromosome 8 (http://www.ncbi.nlm.nih.gov/Homology/).

```
     1                                                          58
Co  AVGKNKRL-KGGKKGAK-KKVVDPFSKKDWYDVKAPAMFNIRNIGKTLVTRTQGTKIA
hu  .......T..................................................
mu  .......T..................................................
ra  .......T..................................................
dr  ......G.S......G..........R.........N..QT.Q......N....QR..
ca  .......IS......G....AA...A......I...SV.SV..V..............
sc  ........S......Q...R.....TR.E.F.I...ST.EN..V.....NKST.L.S.
me  ~~~~~MATARTARSRR.VR..R.KWKE.V..EIY.TPE.GGVF..Y.PANDPS.....
```

```
     59                                                         116
Co  SDGLKGRVFEVSLA-DLQ-D-DED--FRKFKLITEDVQGKNCLTNFHGMDLTRDKM-S
hu  .................N.....VA...............................C.
mu  .................N.....VA...............................C.
ra  .................N.....VA...............................C.
dr  ..YF........VP...K.I.PEHS....R..A....DR.V.C........T..YR.
ca  .E...EH....I......G.....HS...IR.RA..I....V....W...F.T..LR.
sc  ..A.....V..C......G..S..HS...I..RVDE.....L........F.T..LR.
me  ..LVL...A.T...R..TG.PTKHMHRVY..IF..G.T.NKAIAQYY.H.T..EF.K.
```

```
     117                                                        174
Co  MVKKWQT-IEAHVDVKTTDGYLLRLFC-GFTKKRNNQIRKTSYAQH-QVRQIRK-KMM
hu  .......M..................V.....................Q..........
mu  .......M..................V.....................Q..........
ra  .......M..................V.....................Q..........
dr  .......L...I.EA..V......V..I...A.DQQSQ...C..HES...K..SGR.T
ca  L.R...SL...........S.T..M..I......A..QKR.C...SS.I....R...R
sc  ..R....L...N.T...S.D.V..I.AIA..R.QA..VKRH....SSHI.A...V.IS
me  qirrrrsr.d.il....q..hki.tkamvl.ay.x.tkq.s.........d......e
```

```
     175                                                        232
Co  EIMTREVQTNDLK--VNKLIPDSIGKDIEKACQSIYPLHDVFVRKVKMLKKPKFELGK
hu  ............EV............................................
mu  ............EV............................................
ra  ............EV............................................
dr  D.I.N..SGA...QL....AL...A.....S..R.......YI....V....R.DVS.
ca  ...VNQA.SC...DL.Q.F..E...RE....TS.....QN..I....I..A...D...
sc  ..L.K...GST.AQLTS....EV.N.E..N.TKD.F..QNIH.....L..Q...DV.A
me  ..IKAMAKEKTXPQY.QAMLFGEMAEK.KNE.KK.F.IKN.IIY.SEV.SLA.K.ENE
```

```
     233                                        279
Co  LMELHGEGGGSSGKATGDETGAKVERADGYEPPVQESV---------
hu  .........S..............................~~~~~~~~
mu  .............A..........................~~~~~~~~
ra  .............T..........................~~~~~~~~
dr  .L....D...K.VE.VVSSE..PKST.LKVK.......~~~~~~~~~
ca  ...V..D.........YNEDI.T.LD........AE.A.AEPTEVIGA
sc  ..A.....S........E.K.K..T....FKDE.L.T.~~~~~~~~~
me  GFVKEA.EETAEAQE~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

Figure 3. Alignment of seven S3a amino acid sequences. Dots indicate identity to the sequence consensus, dashes indicate gaps to maximize homology. Co, consensus based on aligned sequences; hu, *Homo sapiens* (Metspalu *et al.*, 1992); mu, *Mus musculus* (Rebane *et al.*, 1998); ra, *Rattus norvegicus* (Kho and Zarbl, 1992); dr, *Drosophila melanogaster* (Reynaud *et al.*, 1997); ca, *Catharanthus roseus* (Ito *et al.*, 1991); sc, *Saccharomyces cerevisiae* (Ito *et al.*, 1992); me, *Methanococcus jannaschii* (Swissprot accession nr. P54059).

3.2. Small nucleolar RNAs U73 and U82
(References II and III)

3.2.1. Identification of the snoRNA U73

The presence of *U73* gene copies within the human and mouse *S3a* genes was found during the sequence analysis of these genes. First, the sequence comparison of the human and mouse *S3a* revealed that introns 3 and 4 of both genes contain conserved elements of about 100 bp in length and with 72–85% identity to each other. In addition, box C (UGAUGA) and box D (CUGA) motifs, hallmarks of several snoRNAs, were found in these elements. To test, whether these conserved regions encode a novel snoRNA, Northern blot analysis was performed. Using various antisense oligodeoxynucleotide probes, the mouse *S3a* introns 3 and 4 were demonstrated to encode small stable RNAs of about 73 nt and 69 nt, respectively. In the tested human cell-lines, only the intron 4-encoded RNA of about 65 nt was detected (Fig. 3a in reference II). The length of these RNAs was determined in denaturing polyacrylamide gels and it is accurate within one or two nucleotides. Since snoRNAs up to U72 had been reported earlier (Smith and Steitz, 1997), a novel snoRNA was designated as U73. To distinguish U73 variants specified by different *S3a* introns, the intron 4 encoded U73 was named as U73a and intron 3 encoded as U73b (Fig. 1 in reference II). An alignment of human and mouse U73 encoding sequences is shown in Fig 4a in reference II.

The 5' ends of the U73 snoRNAs were determined by primer extension analysis. As a result, the 5' end for human U73a was mapped to the nt position −6 relative to the box C, for mouse U73a and U73b it was mapped to the position −5 (Fig. 3b and 4a in reference II). In agreement with the results of Northern analysis, no specific primer extension products were detected using oligodoxynucleotides complementary to the human *U73b*.

3.2.2. The function of U73

All the described C/D box snoRNAs contain antisense element(s) of 10–21 nt complementary to pre-rRNA (Bachellerie and Cavaillé, 1998). Since all U73 variants in human and mouse possess features characteristic to C/D box snoRNAs, a potential base pairing between U73 and human pre-rRNA was investigated. As a result, a 12 nt sequence in U73 just upstream the D' box was found to be complementary to a highly conserved region both in the human and mouse 28S rRNAs (including one G-U pairing). Interestingly, the sequence of this antisense element is 100% identical in all expressed variants of *U73*. By contrast, the non-expressed human variant (*U73b*) has a G to A nucleotide substitution in this region that alters the snoRNA–rRNA pairing potential (Fig. 3a in reference II).

In the methylation guide snoRNAs, the fifth nt upstream the D or D' box in the antisense element of the snoRNA invariably faces the methylated nucleotide in the rRNA sequence (Kiss-László *et al.*, 1996, Cavaillé *et al.*, 1996a). The fifth nt upstream the D' box of U73 corresponds to G1739 in human 28S rRNA (nt numbering according to GenBank accession number M11167). None of the determined 2'-O-methyl groups has been placed to this position or to close vicinity of it (Maden, 1990). However, more than 10 2'-O-ribose methylated nucleotides in mammalian 28S rRNA remain to be mapped precisely. Among those undetermined nucleotides there are at least seven methylated G residues occurring in the sequence context G-Gm-G (Maden, 1990). Accordingly, the G1739 is located in the midst of the sequence element GGG as well. Therefore, these data together with the similarity of U73 to other snoRNAs that have been implicated in rRNA ribose methylation strongly suggest that U73 may guide the 2'-O-ribose methylation of the G1739 in human 28S RNA.

3.2.3. snoRNA U82

In vertebrates, the majority of pre-rRNA modification snoRNAs are encoded within the introns of genes encoding the proteins that are involved in nucleolar function, ribosome structure or protein synthesis (Smith and Steitz, 1997, Pelczar and Filipowicz, 1998). In order to find out novel snoRNA-encoding sequences, the introns of orthologous ribosome-associated genes were compared to each other. Interestingly, a highly conserved snoRNA-like region that has no significant similarity (except C, D and D' boxes) to any known snoRNA was found from the 5th intron of human, rat and mouse nucleolin gene (Fig. 1B in reference III). Earlier, it has been shown that mammalian nucleolin gene intron 11 encodes the C/D box snoRNA U20 (Nicoloso *et al.*, 1994). Intron 12 of human, mouse and rat nucleolin gene has been known to specify the H/ACA snoRNA U23 gene (GenBank accession No. AJ007015). As snoRNAs up to U81 had been reported previously (Smith and Steitz, 1998), the conserved snoRNA-like region of the nucleolin gene 5th intron was designated as U82 encoding gene. The structure of the human nucleolin gene and the positions of internally nested snoRNA genes are presented in Fig. 1A in reference III.

To identify the expression of U82 snoRNA, Northern blot analysis of the RNA extracted from HeLa and mouse C127 cells was carried out. Various oligodeoxynucleotides were designed according to human and mouse nucleolin gene sequences and were used as probes (Fig. 1B in reference III). The same amount of the *in vitro* transcript of the cloned human nucleolin gene 5th intron was added to each blot as a stringency reference. As a result, two RNAs about 70 nt and 67 nt, were detected from HeLa and mouse C127 cells with all the oligodeoxynucleotide probes in less stringent hybridization conditions (Fig. 2A and 2B in reference III). The length of these two RNAs was determined in denaturing polyacrylamide gels and it is accurate within one or two nucleotides.

Subsequently, the snoRNA lengths of 70 nt and 67 nt will be used in this study. Rising the stringency of hybridization conditions, only the 70 nt RNA band was visible with one of mouse specific and one of human specific oligodeoxynucleotide probes (Fig. 2A and 2B in reference III), that points to internal sequence differences between the 70 nt and 67 nt RNAs. Using a human U82 specific probe, the *in vitro* synthesized stringency control was always clearly seen together with the 70 nt RNA band even in the conditions, where the 67 nt band disappeared. Therefore, most probably the 70 nt variant of U82 is encoded by nucleolin gene 5th intron in human and mouse cells. The 67 nt variant of U82 could be specified by another gene, genomic locus of which remains unknown.

The 5' ends of the human and mouse U82 snoRNAs were determined by primer extension analysis. As a result, the 5' ends of human and mouse U82 were mapped to the nt position –5 from the box C (Fig. 1B and 2C in reference III).

Since U82 holds characteristic C/D boxes like other snoRNAs containing antisense element(s) complementary to pre-rRNA, a potential base pairing between U82 and human pre-rRNA was investigated. It was found that an 11 nt region upstream the D' box of U82 was complementary to the sequence between 1675 and 1686 nt of human 18S rRNA (nt numbering according to GenBank accession No. X03205). Accordingly, the fifth nucleotide upstream the D' box and within the antisense element of U82 (A1678) has been determined as one of 40 2'-O-methylated nucleotides in human 18S rRNA (Fig 1B in reference III). Thus, probably U82 belongs to the class of snoRNAs that guide site-specific 2'-O-ribose methylation of pre-rRNA and presumably takes part in the methyltransfer reaction of A1678 in human 18S rRNA.

CONCLUSIONS

1. Primary structure of the human rp S3a was determined combining the protein sequence deduced from partial cDNA and direct S3a protein N-terminal sequencing results. The human S3a protein consists of 263 aa, its calculated molecular weight is about 29.8 kDa and it contains unusually large fraction of lysines (38 of 263). Similarly to other rps, the human S3a is basic with a pI of 10.53.

2. The intron-containing genes for the human and mouse rp S3a were cloned and sequenced. The human and mouse *S3a* genes share identical exon/intron structure (are split to 6 exons), exhibit remarkable conservation of the promoter regions (68.8% identity in 250 bp upstream of TSP site) and coding regions (the proteins differ in 2 aa). Human and mouse *S3a* genes display features typical to other rp genes including the CpG-rich 5' region, no consensus TATA or CAAT boxes and TSP within an oligopyrimidine tract. However, the human and mouse *S3a* genes represent a rare subclass of rp genes that possess a long coding sequence in the first exon.

3. Introns 3 and 4 of human and mouse *S3a* were shown to encode a novel snoRNA U73. Both variants of U73 encoding genes are expressed in mouse while only one variant (specified by intron 4) is expressed in tested human cell-lines. U73 contains features similar to intron-encoded antisense snoRNAs that guide site-specific 2'-O-ribose methylation of pre-rRNA. It was proposed that U73 takes part in methylation of the G1739 residue in human 28S rRNA. Human *S3a* and internally encoded *U73* genes were mapped to the human chromosome 4q31.2-3.

4. A novel snoRNA U82 was identified from the human and mouse nucleolin gene in addition to two previously known snoRNAs, U20 and U23. U82 contains features similar to snoRNAs that guide site-specific 2'-O-ribose methylation of pre-rRNA and presumably it takes part in the methyltransfer reaction of A1678 in human 18S rRNA. U82 is present in two variants with estimated length of 70 (± 1) nt and 67 (± 1) nt in HeLa and mouse C127 cells. Probably, the 70 nt variant of U82 is encoded within the 5th intron of nucleolin gene. The 67 nt variant of U82 could be a transcript of another gene, the genomic locus of which remains unknown.

REFERENCES

Abou Elela, S. A., Igel, H. and Ares, M. J. (1996) RNase III cleavages eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell* **85**, 115–124.

Agrawal, M. G. and Bowman, L. H. (1987) Transcriptional and translational regulation of ribosomal protein formation during mouse myoplast differentiation. *J. Biol. Chem.* **262**, 4868–4875.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., Tollervey, D. (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **19**, 5399–5410.

Amaldi, F., Bozzoni, I., Beccari, E. and Pierandrei-Amaldi, P. (1989) Expression of ribosomal protein genes and regulation of ribosome biosynthesis in *Xenopus* development. *Trends Biochem. Sci.* **14**, 175–178.

Amaldi, F. and 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.

Annilo, T. (1998) Studies on mammalian ribosomal protein S7. Dissertation, Tartu University.

Annilo, T., Karis, A., Hoth, S., Rikk, T., Kruppa, J. and Metspalu, A. (1998) Nuclear import and nucleolar accumulation of the human ribosomal protein S7 depends on both a minimal nuclear localization sequence and an adjacent basic region. *Biochem. Biophys. Res. Commun.* **249**, 759–766.

Auclair, D., Lang, B. F., Forest, P. and Desgroseillers, L. (1994) Analysis of genes encoding highly conserved lysine-rich proteins in *Aplysia Californica* and *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **220**, 997–1003.

Avni, D., Biberman., Y. and Meyuhas, O. (1997) The 5' terminal oligopyrimidine tract confers translational control on TOP mRNAs in a cell type- and sequence context-dependent manner. *Nucleic Acid Res.* **25**, 995–1001.

Avni, D., Shama, S., Loreni, F. and Meyuhas, O. (1994) Vertebrate mRNAs with a 5' terminal pyrimidine tract are candidates for translational repression in quiescent cells: characterization of the translational *cis*-regulatory element. *Mol. Cell. Biol.* **14**, 3822–3833.

Bachellerie, J.-P. and Cavaillé J. Small nucleolar RNAs guide the ribose methylations of eukaryotic rRNAs, in: H. Grosjean, R. Benne (Eds.), Modification and editing of RNA, ASM Press, Washington, 1998, pp. 255–272.

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.

Beest, M., Mortin, M. and Clevers, H. (1998) *Drosophila RpS3a*, a novel *Minute* gene situated between the segment polarity genes *cubitus interruptus* and *dTCF*. *Nucleic Acid Res.* **26**, 4471–4475.

Beilharz, T., Beddoe, T., Landl, K., Cartwright, P. and Lithgow, T. (1997) The protein encoded by the *MFT1* gene is a targeting factor for mitochondrial precursor proteins, and not a core ribosomal protein. *FEBS Lett.* **407**, 220–224.

Beltrame, M. and Tollervey, D. (1995) Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. *EMBO J.* **14**, 4350–4356.

Berg, J. M. (1990) Zinc finger domains: hypotheses and current knowledge. *Annu. Rev. Biophys. Chem.* **19**, 405–412.

Bertrand, E., Houser-Scott, F., Kendall, A., Singer, R. H. and Engelke, D. R. (1998) Nucleolar localization of early tRNA processing. *Genes Dev.* **15**, 2463–2468.

Borovjagin, A. V. and Gerbi S. A. (1999) U3 small nucleolar RNA is essential for cleavage at sites 1, 2 and 3 in pre-rRNA and determines which rRNA processing pathway is taken in Xenopus oocytes. *J. Mol. Biol.* **286**, 1347–1363.

Bortolin, M. L., Ganot, P., Kiss, T. (1999) Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. *EMBO J.* **18**, 457–469.

Bortolin, M.-L. and 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.

Bouvet, P., Diaz, J. J. Kindbeiter, K., Madjar, J. J. and Amalric, F. (1998) Nucleolin interacts with several ribosomal proteins through its RGG domain. *J. Biol. Chem.* **24**, 19025–19029.

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

Bulygin, K. N., Matasova, N. B., Graifer, D. M., Veniyaminova, A. G., Yamkovoy, V.I., Stahl, J. and Karpova, G. G. (1997) Protein environment of mRNA at the decoding site of 80S ribosomes from human placenta as revealed from affinity with mRNA analogs--derivatives of oligoribonucleotides. *Biochim. Biophys. Acta* **1351**, 325–332.

Caboche, M. and Bachellerie, J. P. (1977) RNA methylation and control of eukaryotic RNA biosynthesis. Effects of cycloleucine, a specific inhibitor of methylation, on ribosomal RNA maturation. *Eur. J. Biochem.* **74**, 19–29.

Caffarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P. and 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.

Caffarelli, E., Loist, M., Giorgi, C., Fatica, A. and Bozzoni, I. (1998) *In vivo* identification of nuclear factors interacting with the conserved elements of box C/D small nucleolar RNAs. *Mol. Cell. biol.* **18**, 1023–1028.

Caffarelli, E., Maggi, L., Fatica, A., Jiricny, J. and Bozzoni, I. (1997) A novel Mn++-dependent ribonuclease that functions in U16 snoRNA processing in *X. Laevis. Biochem. Biophys. Res. Com.* **233**, 514–517.

Cavaillé, J., and Bachellerie, J. P. (1996c) Processing of fibrillarin-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed common stem-box terminal structure. *Biochimie,* **78**, 443–456.

Cavaillé. J., Hadjiolov, A. A., Bachellerie, J. P. (1996b) Processing of mammalian rRNA precursors at the 3' end of 18S rRNA. Identification of cis-acting signals suggests the involvement of U13 small nucleolar RNA. *Eur. J. Biochem.* **242**, 206–213.

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

Chamberlain, J. R., Pagan-Ramos, Kindelberger, D. W., Engelke, D. R. (1996) An RNase P RNA subunit mutation affects ribosomal RNA processing. *Nucleic Acids Res.* **24**, 3158–3166.

Chamberlain, J. R., Lee, Y., Lane, W. S. and 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.

Chan, Y. L., Olvera, J., Paz, V. and Wool, I. G. (1996) The primary structures of rat ribosomal proteins S3a (the v-*fos* transformation effector) and of S3b. *Biochem. Biophys. Res. Commun.* **228**, 141–147.

Chan, Y. L., Suzuki, K., Olvera, J. and Wool, I. G. (1993) Zinc finger-like motifs in rat ribosomal proteins S27 and S29. *Nucleic Acids Res.* **11**, 649–655.

Chan, Y. L., Suzuki, K. and Wool, I. G. (1995) The carboxyl extensions of two rat ubiquitin fusion proteins are ribosomal proteins S27a and L40. *Biochem. Biophys. Res. Commun.* **215**, 682–690.

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

Chang, D. D. and Clayton, D. A. (1987a) A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. *Science* **235**, 1178–1184.

Chang, D. D. and Clayton, D. A. (1987b) A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. *EMBO J.* 409–417.

Chester, K. A., Robson, L., Begent, R. H. J., Talbot, I. C., Pringle, J. H., Primose, L., Macpherson, A. J. S., Boxer, G., Southall, P. and Malcolm, A. D. B. (1989) Identification of a human ribosomal protein mRNA with increased expression in colorectal tumours. *Biochim. Biophys. Acta* 1009, 297–300.

Chiao, P. J., Shin, D. M., Sacks, P. G., Hong, W. K. and Tainsky, M. A. (1992) Elevated expression of the ribosomal protein S2 gene in human tumours. *Mol. Carcinog.* **5**, 219–231.

Chou, M. M. and Blenis, J. (1995) The 70 kDa S6 kinase: regulation of a kinase with multiple roles in mitogenic signalling. *Curr. Opin. Cell. Biol.* **7**, 806–814.

Chu, S., Archer, R. H., Zengel, J. M. and Lindahl, L. (1994) The RNA of RNase MRP is required for normal processing of ribosomal RNA. *Proc. Natl. Acad. Sci. USA* **91**, 659–663.

Collatz, E., Ulbrich, N., Tsurugi, K., Lightfoot, H. N., MacKinlay, W., Lin, A. and Wool, I. G. (1977) Isolation of eukaryotic ribosomal proteins. *J. Biol. Chem.* 252, 9071–9080.

Colombo, P. and Fried, M. (1992) Functional elements of the ribosomal protein L7a (rpL7a) gene promoter region and their conservation between mammals and birds. *Nucleic Acids Res.* **20**, 3367–3373.

Draptchinskaia, N., Gustavsson, P., Andersson, B., Pettersson, M., Willig, T., Dianzani, I., Ball, S., Tchernia, G., Klar, J., Matsson, H., Tentler, D., Mohandas, N., Carlsson, B. and Dahl, N. (1999) The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nature Gen.* 21, 169–174.

Dunbar, D. A. and Baserga, S. J. (1998) The U14 snoRNA is required for 2'-O-methylation of the pre-18S rRNA in *Xenopus* oocytes. RNA **4**, 195–204.

Dunbar, D. A., Wormsley, S., Agentis, T. M. and Baserga, S. J. (1997) Mpp10p, a U3 small nucleolar ribonucleoprotein component required for pre-18S rRNA processing in yeast. *Mol. Cell. Biol.* **17**, 5813–5812.

van Eenennaam, H., Pruijn, G. J. and van Venrooij, W. J. (1999) hPop 4p: a new protein subunit of the human RNase MRP and RNase P ribonucleoprotein complexes. *Nucleic Acid Res.* **27**, 2465–2472.

Eichler, D. C. and Craig, N. (1994) Processing of eukaryotic ribosomal RNA. *Prog. Nucleic. Acid. Res. Mol. Biol.* **49**, 197–239.

Eisinger, D. P., Dick, F. A. and Trumpower, B. L. (1997) Qsr1p, a 60S ribosomal subunit protein, is required for joining of 40S and 60S subunits. *Mol. Cell. Biol.* **17**, 5136–5145.

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

Feo, S., Davies, B. and Fried, M. (1990) A phage screening method to isolate intron-containing genes in presence of multiple processed pseudogenes. *Nucleic Acids Res.* **16**, 4949–4950.

Fisher, E. M. C., Beer-Romeo, P., Brown, L. G., Ridley, A., McNeil, J. A., Lawrence, J. B., Willard, H. F., Bieber, F. R. and Page, D. C. (1990) Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Tuner syndrome. *Cell* **63**, 1205–1218.

Forster, A. C., Altman, S. (1990) Similar cage-shaped structures for the RNA components of all ribonuclease P and ribonuclease MRP enzymes. *Cell* **10**, 407–409.

Frank, D. N. and Pace, N. R. (1998) Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.* **67**, 153–180.

Frolov, M. V. and Birchler, A. J. (1998) Mutation in P0, a dual function of ribosomal protein/apurinic/apyrimidinic endonuclease, modifies gene expression and position effect variegation in *Drosophila. Genetics* **150**, 1487–1495.

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

Garrett, J. M., Singh, K. K., Vonder Haar, R. A. and Emr, S. D. (1991) Mitochondrial protein import: isolation and characterization of the *Saccharomyces cerevisiae* *MFT1* gene. *Mol. Gen. Genet.* **225**, 483–491.

Ginisty, H., Amalric, F. and Bouvet P. (1998) Nucleolin functions in the first step of ribosomal RNA processing. *EMBO J.* **17**, 1476–1486.

Ginisty, H., Sicard, H., Roger, B. and Bouvet P. (1999) Structure and function of nucleolin. *J. Cell Sci.* **112**, 761–772.

Girard, J. P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D. and Lapeyre, B. (1992) GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. *EMBO J.* **11**, 673–682.

Gordon, M. H., Kucera, G., Salvo, R. and Boss, J. M. (1992) Tumor necrosis factor induces genes involved in inflammation, cellular and tissue repair, and metabolism in murine fibroblasts. *J. Immunol.* **148**, 4021–4027.

Green, R. and Noller, H. F. (1997) Ribosomes and translation. *Annu. Rev. Biochem.* **66**, 679–716.

Hariharan, N., Kelley, D. E. and Perry, R. P. (1989) Equipotent mouse ribosomal protein promoters have similar architecture that includes internal sequence elements. *Genes Dev.* **3**, 1789–1800.

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

Hicks, G. R. and Raikhel, N. V. (1995) Protein import into the nucleus: an integrated view. *Annu. Rev. Cell Dev. Biol.* **11**, 155–188.

Hoth, S. (1993) Zur Biogenese von Ribosomen: S6-genstruktur, S3a-Primärstruktur und -Kernimport.Dissertation, Universität Hamburg.

Hughes, J. M. (1996) Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. *J. Mol. Biol.* Jun 259, 645–654.

Hughes, J. M. and Ares, M. J. (1991) Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *EMBO J.* **10**, 4231–4239.

Ito, M. and Komamine, A. (1993) Precise mapping and molecular characterization of the *MFT1* gene involved in import of a fusion protein into mitochondria in *Saccharomyces cerevisiae. FEBS Lett.* **320**, 125–129.

Ito, M., Sato, T., Fukuda, H. and Komamine, A. (1994) Meristem-specific gene expression directed by the promoter of the S-phase specific gene, *cyc07*, in transgenic *Arabidopsis. Plant Mol. Biol.* **24**, 863–878.

Ito, M., Yasui., A. and Komamine, A. (1992) A gene family homologues to the S-phase specific gene in higher plants is essential for cell proliferation in *Saccharomyces cerevisiae. FEBS Lett.* **301**, 29–33.

Jacobson, M. R., Cao, L. G., Taneja, K., Singer, R. H., Wang, Y. L. and Pederson, T. (1997) Nuclear domains of the RNA subunit of RNase P RNA. *J. Cell Sci.* 110, 829–837.

Jansen, R., Tollervey, D. and Hurt, E. C. (1993) A U3 snoRNP protein with homology to splicing factor PRP4 and G beta domain is required for ribosomal RNA processing. *EMBO J.* **12**, 2549–2558.

Jarrous, N., Eder, P. S., Wesolowski, D. Altman, S. (1999) Rpp14 and Rpp29, two protein subunits of human ribonuclease P. *RNA* **5**, 153–157.

Jefferies, H. B. J., Fumagalli, S., Dennis, P. B., Reinhard, C., Pearson, R. B. and Thomas, G. (1997) Rapamycin suppresses 5' TOP mRNA translation through inhibition of p70s6k. *EMBO J.* **16**, 3693–3704.

Jefferies, H. B. J., Reinhard, C., Kozma, S. C. and Thomas, G. (1994) Rapamycin selectively repress translation of 'polypyrimidine tract' mRNA family. *Proc. Natl. Acad. Sci. USA* **91**, 4441–4445.

Kaspar, R. L., Rychlik, W., White, M. W., Rhoads, R. E. and Morris, D. R. (1990) Simultaneous cytoplasmatic redistribution of ribosomal protein L32 mRNA and phosphorylation of eukaryotic initiation factor 4E after mitogenic stimulation of Swiss 3T3 cells. *J. Biol. Chem.* **265**, 3619–3622.

Kato, S., Sekine, S., Oh, S. W., Kim, N. S., Abe, N., Yokoyama-Kobsyashi, M. and Aoki, T. (1994) Construction of human full-length cDNA bank. *Gene* 150, 243–250.

Kenmochi, N., Kawaguchi, T., Rozen, S., Davis, E., Goodman, N., Hudson, T. J., Tanaka, T. and Page, D. C. (1998) A map of 75 human ribosomal protein genes. *Genome Res.* 8, 509–523.

Kho, C. J., Wang, Y. and Zarbl, H. (1996) Effect of decreased *fte-1* gene expression on protein synthesis, cell growth, and transformation. *Cell Growth. Differ.* **7**, 1157–1166.

Kho, J. K. and Zarbl, H. (1992) *Fte-1*, a v-*fos* transformation effector gene, encodes the mammalian homologue of a yeast gene involved in protein import into mitochondria. *Proc. Natl. Acad. Sci. USA* **89**, 2200–2204.

Kidou, S., Umeda, M., Tsuge, T., Kato, A. and Uchimiya, H. (1994) Isolation and characterization of a rice cDNA similar to the S-phase-specific *cyc07* gene. *Plant Mol. Biol.* **24**, 545–547.

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

Kiss-Lászó, Z., Yves, H., Bachellerie, J.-P., Caizergues-Ferrer, M., Kiss, T. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs, *Cell* **85**, (1996) 1077–1088.

Kiss-Lászó, Z., Yves, H. Kiss, T. (1998) Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J.* **17**, 797–807.

Kodama, I., Ito, M., Hattori, T., Nakamura, K. and Komamine, A. (1991) Isolation of genes that are preferentially expressed at the G_1/S boundary during the cell cycle in synchronized cultures of *Chataranthus roseus* cells. *Plant Physiol.* **95**, 406–411.

Kumar, K. U., Srivastava, S. P. and Kaufman, R. J. (1999) Double-stranded RNA-activated protein kinase (PKR) is negatively regulated by 60S ribosomal subunit protein L18. *Mol. Cell. Biol.* **19**, 1116–1125.

Kuwano, Y., Olvera, J. and Wool, I. G. (1992) The primary structure of rat ribosomal protein S5. A rat ribosomal protein present in the rat genome in a single copy. *J. Biol. Chem.* **267**, 25304–25308.

Lafontaine, D. L., Bousquet-Antonelli, C., Henry, Y., Caisergues-Ferrer, M., and Tollervey, D. (1998) The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes. Dev.* **12**, 527–537.

Lafontaine, D. L., Tollervey, D. (1999) Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA* **5**, 455–467.

Leader, D. J., Clark, G. P., Watters, J., Beven, A. F., Shaw, P. J. and 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.

Lecomte, F., Szpirer, J. and Szpirer, C. (1997) The S3a ribosomal protein gene is identical to the *Fte-1* (v-*fos* transformation effector) gene and the TNF-α-induced *TU-11* gene, and its transcript level is altered in transformed and tumor cells. *Gene* **186**, 271–277.

Lee, S. J. and Baserga, S. J. (1999) Imp3p and imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. *Mol. Cell Biol.* **19**, 5441–5452.

Levy, S., Avni, D., Hariharan, N., Perry, R. P. and Meyuhas, O. (1991) Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control. *Proc. Natl. Acad. Sci. USA* **88**, 3319–3323.

Liang, W. Q. and Fournier, M. J. (1995) U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing. *Genes. Dev.* **9**, 2433–2443.

Louie, D. F., Resing, K. A. Lewis, T. S. and Ahn, N. G. (1996) Mass spectrometric analysis of 40S ribosomal proteins from Rat-1 fibroplasts. *J. Biol. Chem.* **271**, 28189–28198.

Lowe, T. M. and Eddy, S. R. (1999) Computational screen for methylation guide snoRNAs in yeast. *Science* **283**, 1168–1171.

Lutsch, G., Stahl, J., Kärgel., H. J., Noll, F. and Bielka, H. (1990) Immunoelectron microscopic studies on the location of ribosomal proteins on the surface of 40S ribosomal subunit from rat liver. *Eur. J. Cell Biol.* **51**, 140–150.

Lygerou, Z., Allmang, C., Tollervey, D. and Seraphin, B. (1996) Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP *in vitro. Science.* **272**, 268–270.

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

Mager, W. H. and Planta, R. J. (1990) Multifunctional DNA-binding proteins mediate transcription activation of yeast ribosomal protein genes. *Biochim. Biophys. Acta* **1050,** 351–355.

Mager, W. H., Planta, R. J., Ballesta, J.-P. G., Lee, J. C., Mizuta, K., Suzuki, K., Warner, J. R. and Woolford, J. (1997) A new nomenclature for the cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae. Nucleic. Acids. Res.* **25**, 4872–4875.

Maxwell, E. S. and Fournier, M. J. (1995) The small nucleolar RNAs, *Ann. Rev. Biochem.* **35**, 897–934.

Meyuhas, O., Thompson, E. A. J. and Perry. R. P. (1987) Clycocorticoids selectively inhibit translation of ribosomal protein mRNAs in P1798 lymphosarcoma cells. *Mol. Cell. Biol.* **7**, 2691–2699.

Mazuruk, K., Schoen, T. J., Chader, G. J. Iwata, T. and Rodriguez, I. R. (1996) Structural organization and chromosomal localization of the human ribosomal protein L9 gene. *Biochim. Biophys. Acta* **1305**, 151–162.

Michot, B., Joseph, N., Mazan, S. and Bachellerie, J. P. (1999) Evolutionarily conserved structural features in the ITS2 of mammalian pre-rRNAs and potential interactions with the snoRNA U8 detected by comparative analysis of new mouse sequences. *Nucleic Acids Res.* **27**, 2271–2282.

Mishra, K. R. and Eliceiri, G. L. (1997) Three small nucleolar RNAs that are involved in ribosomal RNA precursor processing. *Proc. Natl. Acad. Sci. USA* **94**, 4972–4977.

Morrissey, J. P. and Tollervey, D. (1993) Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol. Cell. Biol.* **13**, 2469–2477.

Morrissey, J. P. and Tollervey, D. (1997) U14 small nucleolar RNA makes multiple contacts with the pre-ribosomal RNA. *Chromosoma* **105**, 515–522.

Mount, S. M. (1982) A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**, 459–472.

Naora, H. (1999) Involvement of ribosomal proteins in regulation cell growth and apoptosis: translational modulation or recruitment for extraribosomal activity *Immunol. Cell Biol.* **77**. 197–205.

Naora, H., Nishida, Y., Adachi, M. and Naora, H. (1995) Association of *nbl* gene expression and glycocorticoid-induced apoptosis in mouse thymus *in vivo. Immunology* **85**, 63–68.

Naora, H., Takai, I., Adachi, M. and Naora, H. (1998) Altered cellular responses by varying expression of a ribosomal protein gene: sequential coordination of enhancement and suppression of ribosomal protein S3a gene expression induces apoptosis. *J. Cell Biol.* **141**, 741–753.

Neumann, F., Hemmerich, P., von Mikecz, A., Peter, H. H. and Krawinkel, U. (1995) Human ribosomal protein L7 inhibits cell-free translation in reticulocyte lysates and affects the expression of nuclear proteins upon stable transfection into Jurkat T-lymphoma cells. *Nucleic Acids Res.* **23**, 195–202.

Neumann, F. and Krawinkel, U. (1997) Constitutive expression of human ribosomal protein L7 arrests the cell cycle in G1 and induces apoptosis in jurkat T-lymphoma cells. *Exp. Cell Res.* **230**, 252–261.

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

Nicoloso, M., Caizergues-Ferrer, M., Michot, B., Azum, M.-C., Bachellerie, J.-P. (1994) U20, a novel small nucleolar RNA, is encoded in an intron of the nucleolin gene in mammals. *Mol. Cell. Biol.* **14**, 5766–5776.

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

Nolte, D., Taimor, G., Kalf-Suske, M. and Seifart and K.H. (1996) The human S3a ribosomal protein: sequence, localisation and cell-free transcription of the functional gene. *Gene* **169**, 179–185.

Noon, K. R., Bruenger, E. and McCloskey, J. A. (1998) Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hypertermophile *Sulfolobus solfataricus. J. Bacteriol.* **180**, 2883–2888.

Ofengand, J. and Fournier, M. J. The pseudouridine residues of rRNA: number, location, biosynthesis, and function, in: H. Grosjean, R. Benne (Eds.), Modification and editing of RNA, ASM Press, Washington, 1998, pp. 229–254.

Ofengand, J. and 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.

Olvera, J. and Wool, I. G. (1993) The carboxyl extension of a ubiquitin-like protein is rat ribosomal protein S30. *J. Biol. Chem.* **268**, 17967–17974.

Ouarzane, M., Labbe, M. and Pery, P. (1998) *Eimeria tenella*: cloning and characterization of cDNA encoding a S3a ribosomal protein. *Gene* **225**, 125–130.

Parker, K. A. and Steitz, J. A. (1987) Structural analysis of the human U3 ribonucleoprotein particle reveals a conserved sequence available for base pairing with pre-rRNA. *Mol. Cell. Biol.* **7**, 2899–2913.

Pata, I. and Metspalu, A. Structural characterization of the mouse ribosomal protein S6-encoding gene. *Gene* **175**, 241–245.

Peculis, B. A. (1997) The sequence of the 5' end of the U8 small nucleolar RNA is critical for 5.8S and 28S rRNA maturation. *Mol. Cell. Biol.* **17**, 3702–3713.

Peculis, B. A. and Steitz, J. A. (1993) Disruption of U8 small nucleolar snRNA inhibits 5.8S and 28S rRNA processing in the *Xenopus* oocyte. *Cell* **73**, 1233–1245.

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

Pellizoni, L., Lotti, F., Rutjes, S. A. and Pierandrei-Amaldi, P. (1998) Involvement of the *Xenopus laevis* Ro60 autoantigen in the alternative interaction of La and CNBP proteins with 5' UTR of L4 ribosomal protein mRNA. *J. Mol. Biol.* **281**, 593–608.

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

Philipsen, S. and Suske, G. (1999) A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. *Nucleic Acids Res.* **27**, 2991–3000.

Planta, R. J. (1997) Regulation of ribosome synthesis in yeast. *Yeast* **13**, 1505–1518.

Pluk, H., Soffner, J., Lührmann, R. and van Venrooij, W. J. (1998) cDNA cloning and characterization of the human U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. *Mol. Cell. Biol.* **18**, 488–498.

Pluk, H., van Eenennaam, H., Rutjes, S. A., Pruijn, G. J. and van Venrooij, W. J. (1999) RNA-protein interactions in the human RNase MRP ribonucleoprotein complex. *RNA* **5,** 512–524.

Pogue-Geile, K., Geiser J. R., Shu, M., Miller, C., Wool, I. G., Meisler, I. and Pipas, J. M. (1991) Ribosomal protein genes are overexpressed in colorectal cancer: isolation of a cDNA clone encoding the human S3 ribosomal protein. *Mol. Cell. Biol.* **11**, 3842–3849.

Pugh, B. F. and Tjian, R. (1990) Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* **61**, 1187–1197.

Ramakrishnan, V. and White, W. (1998) Ribosomal protein structures: insights into the architecture, machinery and evolution of the ribosome. *Trends Biol. Sci.* **6**, 208–212.

Reddy, R. and Shimba, S. (1995) Structural and functional similarities between MRP and RNase P. *Mol. Biol. Rep.* **22**, 81–85.

Remacha, M., Jimenes-Diaz, A., Santos, C., Briones, E., Zambrano, R., Rodriguez Gabriel, M. A., Guarinos, E. and Ballesta, J. P. (1995) Proteins P1, P2, and P0, components of eukaryotic ribosome stalk. New structural and functional aspects. *Biochem. Cell Biol.* **73**, 959–968.

Reynaud. E., Bolshakov, V.N., Barajas, V., Kafatos, F. C. and Zurita, M. (1997) Antisense suppression of the putative ribosomal protein S3A gene disrupts ovarian development in *Drosophila melanogaster. Mol. Gen. Genet.* **256**, 462–467.

Sæbøe-Larssen, S., Lyamouri, M., Merriam, J., Oksvold, M. P. and Lambertsson, A. (1997) Ribosomal protein insufficiency and the minute syndrome in *Drosophila*: a dose-response relationship. *Genetics* **148**, 1215–1224.

Samarsky, D. A., Schneider, G. S. and Fournier, M. J. (1996) An essential domain in *Saccharomyces cerevisiae* U14 snoRNA is absent in vertebrates, but conserved in other yeast. *Nucleic. Acids. Res.* **24**, 2059–2066.

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

Schmitt, M. E. and Clayton, D. A. (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae. Mol. Cell. Biol.* **13**, 7935–7941.

Schmitt, M. E. and Clayton, D. A. (1994) Characterization of a unique protein component of yeast RNase MRP: An RNA-binding protein with a zinc-cluster domain. *Genes Dev.* **8**, 2617–2628.

Segal, D. M. and Eichler, D. C. (1991) A nucleolar 2'-O-methyltransferase. Specificity and evidence for its role in the methylation of mouse 28S precursor ribosomal RNA. *J. Biol. Chem.* **266**, 24385–24389.

Sharma, K. and Tollervey, D. (1999) Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA isrequired for pre-rRNA processing. *Mol. Cell. Biol.* **19**, 6012–6019.

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

Smith, C. M. and Steitz, J. A. (1998) Classification of *gas5* as a multi-small-nucleolar-RNA (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.

Smith, J. E., Cooperman, B. S. and Mitchell, P. (1992) Methylation sites in *E. Coli* ribosomal RNA: localization and identification of four new sites of methylation in 23S rRNA. *Biochemistry* **31**, 10825–10834.

Starkey, C. R, Menon, R. P., Prbhu, S. and Levy, L. S. (1996) Primary sequence and evolutional conservation of ribosomal protein genes from the domestic cat. *Biochem. Biophys. Res. Com.* **220**, 648–652.

Starkey, C. R. and Levy, L. S. (1995) Identification of differentially expressed genes in T-lymphoid malignancies in an animal model system. *Int. J. Cancer* **62**, 325–331.

Stewart, M. J. and Thomas, G. (1994) Mitogenesis and protein synthesis: a role for ribosomal protein S6 phosphorylation? *BioEssays* **16**, 809–815.

Suzuki, K., Olvera, J. and Wool, I.G. (1991) Primary structure of rat ribosomal protein S2. A ribosomal protein with arginine-glycine tandem repeats and RGGF motifs that are associated with nucleolar localization and binding to ribonucleic acids. *J. Biol. Chem.* **266**, 20007–20010.

Swoboda, K. R., Broadbent, I. D., Bertram, G., Budge, S., Gooday, G. W., Gow, N. A. R. and Brown, A. J. B. (1995) Structure and regulation of a *Candida albicans RP10* gene which encodes an immunogenic protein homologous to *Saccharomyces cerevisiae* ribosomal protein 10. *J. Bacteriol.* **177**, 1239–1236.

Todorov, T., Noll, F. and Hadjiolov, A. A. (1983) The sequential addition of ribosomal proteins during formation of the small ribosomal subunit in friend erytroleukemia cells. *Eur. J. Biochem.* **131**, 271–275.

Tollervey, D. (1987) A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *EMBO J.* **6**, 4169–4175.

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

Tollervey, D., Lehtonen, H., Carmo-Fonseca, M. and Hurt, E. C. (1991) The small nucleolar RNP protein NOP1 (fibrillarin) is required pre-rRNA processing in yeast. *EMBO J.* **10**, 573–583.

Tollervey, D., Lehtonen, H., Jansen, R., Kern, H. and Hurt, E. C. (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* **72**, 443–457.

Tycowski, K. T., Shu, M.-D. and Steitz, J. A. (1996) A mammalian gene with introns instead of exons generating stable RNA products. *Nature* **379**, 3158–3166.

Tycowski, K. T., Shu, M.D. and Steitz, J. A. (1994) Requirement for intron-encoded U22 small nucleolar RNA in 18S ribosomal RNA maturation. *Science* **266**, 1558–1561.

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

Vaarala, M. H., Porvar, K. S., Kyllönen, A. P, Mustonen, M. V. J., Lukkarinen, O. and Vihko, P. T. (1998) Several genes encoding ribosomal proteins are over-expressed in prostate-cancer cell lines: confirmation of L7a and L37 over-expression in prostate-cancer tissue samples. *Int. J. Cancer* **78**, 27–32.

Vanin, E. F. (1985) Processed pseudogenes: characteristics and evolution. *Annu. Rev. Genet.* **19**, 253–272.

47

Wagner, M. and Perry, R. P. (1985) Characterization of the multigene family encoding the mouse S16 ribosomal protein: strategy for distinguishing an expressed gene from its processed pseudogene counterparts by an analysis of total genomic DNA. *Mol. and Cell. Biol.* **5**, 3560–3576.

Warner, J. R. and Nierras, C. R. (1998) Trapping human ribosomal protein genes. *Genome Res.* **8**, 419–421.

Watkins, N. J, Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Luhrmann, R. (1998a) 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.

Watkins, N. J, Newman, D. R., Kuhn, J. F. and Maxwell, E. S. (1998b) *In vitro* assembly of the mouse U14 snoRNP core complex and identification of a 65-kDa box C/D-binding protein. *RNA* **4**, 582–593.

Watkins, N. J., Leverette, R. D, Xia, L., Andrews, M.T, Maxwell, E.S. (1996) Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *RNA* **2**, 118–133.

Westermann, P. and Nygard, O. (1983) The spatial arrangement of the complex between eukaryotic initiation factor eIF-3 and 40 S ribosomal subunit. Cross-linking between factor and ribosomal protein. *Bichim. Biophys. Acta* **741**, 103–108.

Williamson, N. A., Raliegh, J., Morrice, N. A. and Wettenhall, R. E. (1997) Post-translational processing of rat ribosomal proteins. Ubiquitous methylation of Lys22 within the zinc-finger motif of RL40 (carboxy-terminal extension protein 52) and tissue-specific methylation of Lys4 in RL29. *Eur. J. Biochem.* **246**, 786–793.

Wool, I. G. (1996) Extraribosomal functions of ribosomal proteins. *Trends Biochem. Sci.* **21**, 164–165.

Wool, I. G. Chan, Chan, Y. L. and Glück, A. (1995) Structure and evolution of mammalian ribosomal proteins. *Biochem. Cell. Biol.* **73**, 933–947.

Wu P., Brockenbrough, J. S., Metcalfe, A. C., Chen, S. and Aris, J. P. (1998) Nop5p is a small nucleolar ribonucleoprotein component required for pre-18S rRNA processing in yeast. *J. Biol. Chem.* **273**, 16453–16463.

Yacoub, A., Augeri, L., Kelley, M. R., Doetsch, P. W. and Deutsch, W. A. (1996a) A *Drosophila* ribosomal protein contains 8-oxoguanine and a basic site DNA repair activities. *EMBO J.* **15**, 2306–2312.

Yacoub, A., Kelley, M. R., and Deutsch, W. A. (1996b) *Drosophila* ribosomal protein P0 contains apurinic/apyrimidinic endonuclease activity. *Nucleic Acids Res.* **24**, 4298–4303.

Yoganathan, T. Horikoshi, M., Roeder, R. G. and Sells, B. H. (1993) Direct binding of yeast transcription factor (TFIID) to the ribosomal protein (rpL32) TATA-less promoter sequence. *FEBS Lett.* **326**, 163–166.

Yu, Y. T., Scharl, E. C., Smith, C. M. and Steitz, J. A. (1999) The growing world of small ribonucleoproteins. The RNA world, Second edition, Cold Spring Harbor laboratory press. pp. 487–524.

Zanchin, T. I. and Goldfarb, D. S. (1999) The exosome subunit Rrp43p is required for the efficient maturation of 5.8S, 18S and 25S rRNA. *Nucleic Acids Res.* **27**, 1283–1288.

HIIRE JA INIMESE RIBOSOOMIVALGU S3a GEENIDE NING INTRON-KODEERITAVATE VÄIKESTE TUUMAKESE RNAde U73 JA U82 ISELOOMUSTAMINE

Kokkuvõte

Eukarüootsetes rakkudes toimub tsütoplaasmaatiliste ribosoomide biogenees tuumakestes. Ribosomaalne DNA transkribeeritakse pika pre-rRNA molekulina otse tuumakeses, ribosoomivalgud ja 5S rRNA transporditakse sinna mujalt rakust. Ribosoomide biogeneesi käigus pre-rRNAd töödeldakse ning tulemusena tekkivad 18S, 5.8S ja 25S või 28S rRNA assambleeritakse koos 5S rRNA ja ribosoomivalkudega ribosoomi suureks ja väikeseks allühikuks. Samaaegselt rDNA transkriptsiooniga või vahetult pärast seda paljud rRNA nukleotiidid modifitseeritakse. Peamiselt leiavad aset riboosijääkide 2'-O-metüülimine ja uridiinide muutmine pseudouridiinideks (ülevaade Eichler ja Craig, 1995). Viimaste aastate jooksul on ilmnenud, et pre-rRNA modifitseerimisel, töötlemisel ja pakkimisel ribosoomi osaleb kõrvuti valkudega suur hulk väikesi tuumakese RNAsid (snoRNAsid) (ülevaade Smith ja Steitz, 1997).

Peaaegu kõik selgroogsete rRNA modifitseerimisel osalevad snoRNAd transkribeeritakse mitte iseseisvate geenide poolt, vaid valke kodeerivate geenide intronites. Hiljuti leiti, et selgroogsetes organismides kuuluvad kõik seni kirjeldatud intron-kodeeritavate snoRNAde peremeesgeenid 5' TOP geenide perekonda (Smith ja Steitz, 1998). Väärib märkimist, et 5' TOP geeniperekonna moodustavad peamiselt ribosoomi biogeneesi ning funktsioneerimisega seotud valkude, sealhulgas ribosoomivalkude, geenid (ülevaade Amaldi and Pierandrei-Amaldi, 1997). Võib oletada, et selline snoRNA geenide paiknemine 5' TOP geenide intronites võimaldab snoRNAde ja ribosoomiga seotud geenide ekspressiooni koosregulatsiooni transkriptsiooni tasemel ning toetab translatsiooniaparaadi komponentide proportsionaalset laekumist.

Ribosomaalsete mutatsioonide mõju ning võimalikku seost haiguste tekkega inimestel on vähe uuritud. Arvestades valgu sünteesi fundamentaalset rolli rakus, on ennustatud, et translatsiooniga seotud valkude geenide defektid põhjustavad inimesel ja teistel imetajatel alati varase embrüonaalse surma (ülevaade Warner ja Nierras, 1998). On siiski võimalik, et translatsiooniaparaadiga seotud geenide mutatsioonid avalduvad haigustena. Näiteks avastati hiljuti, et ribosoomivalgu S19 geeni mutatsioonid võivad põhjustada Diamond-Blackfani aneemiat (Drapinskaia et al., 1999).

Üheks vältimatuks etapiks ribosoomivalkude ja teiste translatsioonis ning ribosoomi biogeneesis osalevate komponentide funktsiooni uurimisel on nende

geenide iseloomustamine. Käesolevas töös on käsitletud ribosoomivalku S3a ning teda kodeerivaid hiire ja inimese geene. Et hiire ja inimese *S3a* geeni intronitest leiti seni kirjeldamata snoRNA järjestus, on uurimuse teine pool on seotud snoRNAdega.

Töö esimeses etapis eraldati ja klooniti inimese ribosoomivalgu S3a cDNAd. Pikima cDNA järgi tuletatud valgu järjestuse ja otsesest N-terminaalsest valgu sekveneerimisest saadud järjestuse kombineerimisel kirjeldati inimese S3a primaarstruktuur. Inimese S3a koosneb 263 aminohappest, tema arvutuslik molekulmass on 29,8 kDa. Sarnaselt teistele ribosoomivalkudele on S3a aluseline valk (pI on 10,53).

Järgnevalt klooniti ja sekveneeriti inimese ja hiire ribosoomvalgu S3a geenid. Leiti, et inimese ja hiire *S3a* geenidel on sarnane ekson-intron struktuur: mõlematel geenidel on kuus eksonit ja viis intronit, mis paiknevad identsetel positsioonidel. Inimese ja hiire *S3a* geenid omavad sarnaseid jooni teiste ribosoomivalgu geenidega — CpG-rikas 5'-piirkond, TATA või CAAT järjestuse puudumine promootorpiirkonnas ning transkriptsiooni alguskoha paiknemine oligopürimidiinses alas.

Inimese ja hiire *S3a* geenide järjestuse võrdlemisel ilmnes, et mõlemate geenide intronid 3 ja 4 sisaldavad uudset snoRNAd, U73 kodeerivaid järjestusi. *Northern*-analüüs näitas, et hiirel on mõlemad U73 geenid aktiivsed, kuid inimese uuritud rakuliinides ekspresseeritakse tõenäoliselt ainult ühte, 4. intronis sisalduvat U73 varianti. Võrdlusest teiste snoRNAdega ilmnes, et U73-l on samasugused iseloomulikud jooned nagu intron-kodeeritavatel snoRNAdel, mis osalevad pre-rRNA 2'-O-metüülimisel. Tõenäoliselt võtab U73 osa inimese 28S rRNA 2'-O-metüülimisest positsioonis G1739. Inimese *S3a* geen ja U73 kodeerivad järjestused leiti paiknevat kromosoomis 4, piirkonnas q31.2-3.

Lisaks ülal kirjeldatule analüüsiti ribosoomiga seotud ning paralleelselt eri liikides sekveneeritud geenide intronite järjestusi. Seni kirjeldamata snoRNA, U82, leiti inimese ja hiire nukleoliini geeni viiendast intronist. Varem oli teada, et nukleoliini geeni intronid 11 ja 12 kodeerivad vastavalt snoRNAsid U20 ja U23. U82 geeni ekspressiooni näidati *Northern*-analüüsiga, kusjuures selgus, et U82 esineb hiire ja inimese rakkudes kahesuguse pikkusega (umbes 70 nt ja 67 nt) RNAna. Tõenäoliselt kodeeritakse U82 70 nt pikkust varianti nukleoliini geeni poolt, 67 nt pikkune U82 variant võib pärineda mõnest teisest, seni kirjeldamata lookusest. Oletatavalt on U82 vajalik inimese 18S rRNA 2'-O-metüülimisel positsioonis A1678.

ACKNOWLEDGEMENTS

PUBLICATIONS

Metspalu, A., Rebane, A., Hoth, S., Pooga, M., Stahl, J. and Kruppa, J. (1992)
Human ribosomal protein S3a: cloning of the cDNA and
primary structure of the protein.
Gene 119, 313–316.

Human ribosomal protein S3a: cloning of the cDNA and primary structure of the protein

(cDNA expression library screening; cDNA analysis by PCR; mRNA size; number of genes; N-terminal amino acids)

Andres Metspalu[a], Ana Rebane[a], Stefan Hoth[b], Margus Pooga[a], Joachim Stahl[c] and Joachim Kruppa[b]

[a] *Estonian Biocenter, Tartu, Estonia;* [b] *University of Hamburg, Department of Molecular Biology, Grindelallee, Hamburg, Germany. Tel. (49-40)41235683; and* [c] *Max Delbrück Center for Molecular Medicine, Berlin-Buch, Germany. Tel. (49-372)9463785*

SUMMARY

The amino acid (aa) sequence of human ribosomal protein S3a (hRPS3a) was deduced partially from the nucleotide sequence of the corresponding cDNA and confirmed by direct aa sequencing from the N terminus of the purified hRPS3a protein. The cDNA clone was isolated from a cDNA expression library in the pEX vector using antibodies. The hRPS3a protein has 263 aa and its calculated M_r is 29 813.

INTRODUCTION

After the pioneering work from Perry's laboratory (Meyuhas and Perry, 1980), several groups have cloned and sequenced mammalian r-protein–encoding cDNAs and sequenced proteins. All these data have been reviewed recently (Wool et al., 1990). However, only a few human r-protein (hRP) sequences or cDNAs are known. These are

Correspondence to: Dr. A. Metspalu, Laboratory of Gene Expression, Estonian Biocenter, 2 Jakobi St., EE2400 Tartu, Estonia.
Tel. (7-01434)35211; Fax (7-01434)35430;
E-mail: andres@ebc.tartu.cw.su.

Abbreviations: aa, amino acid(s); BSA, bovine serum albumin; bp, base pair(s); cDNA, DNA complementary to RNA; 2D, two-dimensional; h, human; HPLC, high-performance liquid chromatography; *hRP S3a*, gene (DNA) encoding hRPS3a; kb, kilobase(s) or 1000 bp; LB, Luria–Bertani (medium); nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; PAGE, polyacrylamide-gel electrophoresis; PBS, phosphate-buffered saline (0.15 M NaCl/0.15 M Na₃·phosphate pH 7.4); PCR, polymerase chain reaction; PTH, phenylthiohydantoin; r, ribosomal; r-protein, ribosomal protein; RPS3a, r-protein S3a; SDS, sodium dodecyl sulfate; SM buffer, 0.1 M NaCl/0.015 M MgSO₄·H₂O/0.05 M Tris·HCl pH 7.5/0.01% gelatin.

hRPS6 (Wettenhall et al., 1988; Lott and Mackie, 1988; Heinze et al., 1988), P_o, P_1, P_2 (Rich and Steitz, 1987), L31 (Nobori et al., 1989), L35a (Herzog et al., 1990), S3 (Zhang et al., 1990), and S4 (Fischer et al., 1990). The situation with human r-protein-encoding genes is even more difficult only two genes *S14* and *S17* have been isolated and characterized (Rhoads et al., 1986; Chen and Roufa, 1988). The insufficiency of the structural information about the human ribosomes is obvious, since there are approx. 85 different r-proteins in the whole ribosome. The complete structure of r-proteins and their genes is a prerequisite for asking relevant biological questions about their biological regulation and function. We report here the primary structure of hRPS3a.

EXPERIMENTAL AND DISCUSSION

(a) Isolation and determination of the nt sequence of the cDNA for hRPS3a

A human liver cDNA expression library in pEX vectors (Stanley and Luzio, 1984) was used for screening with a rat r-protein S3a-specific antibody. Antibodies have been

characterized elsewhere (Lutsch et al., 1990). Colony screening was done essentially as in Stanley and Luzio (1984). Briefly, five large Nunc plates (22 × 22 cm), 15 000 colonies per plate were incubated 12 h at 30°C on LB media with ampicillin (100 μg/ml). After colony lifting, nitrocellulose filters were incubated at 42°C for 2 h and colonies were lysed by incubating filters in an oven for 15 min at 95°C. Proteins were transferred onto the membrane according to Western blot protocol for 30 min at 50 V. Filters were treated with DNase (10 μg/ml), washed (1% BSA/0.1% Triton X-100 in PBS) and incubated with the antibody solution (2 μg/ml) for 60 min. After washing, [^{125}I]Protein A (IM.144 Amersham) was added (5 × 10^5 cpm/ml). Primary screening resulted in 16 positive colonies. After two additional rounds of colony purification the plasmid DNA was isolated and the longest insert was excised with PstI and recloned into M13mp8 and M13mp19 vectors and sequenced. The insert was 721 bp long and could not encode the entire S3a protein. We rescreened the same library with radioactive probes, but could not find a longer insert. We plated out a new human cDNA library in λZAPII vector (Stratagene, La Jolla, CA, Cat. No. 936201) according to the supplier's instructions. To speed up the screening procedure we used a new approach based on PCR. First screening of plaques was performed with a ^{32}P-labeled cDNA fragment, generated by PCR. Then agar plugs from positive areas were incubated in 1 ml SM buffer (Sambrook et al., 1989) for 1 h and 10 μl of that material was used directly in a subsequent amplification reaction. The cDNA-specific oligo primer (21 mer, complementary to the nt 400–420 on Fig. 1) was synthesized, and the T3 or T7 17-mer primer (Stratagene) was used as a second primer in two different sets of reactions. In this way only the 5′ part of the S3a cDNA was amplified regardless of the orientation of insert and the length of the PCR product is reflecting the length of the cDNA insert. This is a fast, simple and convenient method for screening the missing 5′ parts of cDNAs. However, in our case the longest insert came from a cloning artifact, during the preparation of the library the two cDNA-s were ligated together via their 5′ ends. The second longest clone gave us the additional 115 nt to the existing clone, but this library did not contain a full-length hRPS3a cDNA. It is possible that some mammalian r-protein mRNAs form a strong secondary structure around the AUG start codon, which is difficult for the reverse transcriptase to read through. In many cases only a few nt have been found before the first AUG codon, e.g., for hRPS6 (Lott and Mackie, 1988), rat RPL9 (Suzuki et al., 1990), rat L7a (Nakamura et al., 1989), human L31 (Nobari et al., 1989), rat L35a (Tanaka et al., 1986), rat S11 (Tanaka et al., 1985), rat L18 (Devi et al., 1988). Altogether we isolated a cDNA clone with an 836-bp insert having a single ORF of 262 aa, one UAA stop codon, a

Fig. 1. The hRP S3a sequence. (a) N-terminal aa sequence of the hRPS3a protein determined directly. (b) The cDNA sequence of the hRP S3a and deduced aa sequence. cDNA was sequenced from both strands using Sanger's dideoxy technique with [^{35}S]dATP and T7 DNA polymerase. The cDNA sequence will appear in the EMBL/GenBank Nucleotide Sequence Databases under the accession No. M 77234.

3′-noncoding region of 43 nt and a poly(A) tail (Fig. 1b). However, this fragment did not contain the complete 5′ terminus.

(b) N-terminal aa sequence analysis and the structure of hRPS3a

In order to determine the N-terminal aa sequence and to confirm the cDNA clone, we sequenced directly the purified hRPS3a. The analysis was performed using a 477A Applied Biosynthesis Protein-Peptide Sequencing System. Phenylthiohydantoin(PTH)-aa were identified on-line with a 120A Applied Biosystems PTH Analyzer by reverse-phase HPLC using a Brownlee PTH-C18 cartridge (2.1 × 220 mm). All sequencer chemicals and supplies were purchased from Applied Biosystems (Foster City, CA). hRPS3a was eluted from 2D-PAGE plugs (Lastic and McConcey, 1976), loaded onto SDS–PAGE (Laemmli, 1970) and electroblotted onto a ProBlott membrane (ABI User Bulletin, 1988). The protein band (2 × 55 mm) was excised, washed with methanol, rinsed twice in Milli Q water and air dried. This strip was cut into smaller fragments (2 × 3 mm), placed on Biobrene-treated filters and 10 aa were determined. The result was: Ala-Val-Gly-Lys-Asn-Lys-Arg-Leu-Thr-Lys (see Fig. 1a). As it turned out, our cDNA missed only four nt from the coding part, including the

AUG start codon. Since the protein sequence started with Ala, the first Met must have been removed during protein processing. The next 9 aa were exactly the same as had been deduced from the cDNA sequence (Fig. 1a). Human r-protein S3a consists of 263 aa and its M_r is 29 813. Homology analysis in EMBL DNA data bank, Swiss Prot and NBRF/PIR protein data bank did show a yeast sequence, 65% homologous to hRPS3a protein just downstream from *SIR3* gene (Shore et al., 1984).

(c) Size of the mRNA for hRPS3a

Ten μg poly(A)$^+$ RNA from HeLa cells was separated by electrophoresis in 1.2% agarose gel in the presence of 2.2 M formaldehyde (Sambrook et al., 1989). After hybridization with hRPS3a-specific ^{32}P-labeled riboprobe from pGEM4Z one band of about 900 nt was detected (Fig. 2).

(d) The copy number of the *hRP S3a* gene

Chromosomal DNA was isolated from HeLa cells and digested with *Bam*HI, *Pst*I, *Kpn*I, *Sac*I, *Bam*HI+*Hind*III and *Hind*III. With labeled *hRP S3a* cDNA we could see 6–8 bands (Fig. 3), suggesting that as for all other mammalian r-protein-encoding genes, *hRP S3a* belongs to the large gene family.
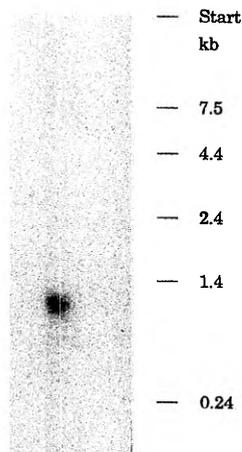

Fig. 3. Southern blotting of HeLa DNA and probing with ^{32}P-labeled *hRP S3a* cDNA. Lanes: 1, *Bam*HI; 2, *Hind*III; 3, *Kpn*I; 4, *Pst*I; 5, *Sac*I; 6, *Bam*HI+*Hind*III. Marker are the *Eco*RI+*Hind*III digest of λ DNA (see right margin). 1% agarose gel was used.

Fig. 2. Northern hybridization of the HeLa cell poly(A)$^+$ RNA with *hRP S3a* probe. The size of *hRP S3a* mRNA is approx. 0.9 kb (on 1.2% agarose gel). Size markers for *hRP S3a* was a 0.24–9.5 kb RNA ladder (see right margin). Hybridization was performed without formamide in 6 × SSC/5 × Denhardt's/0.1% SDS/0.5 × 10^6 cpm per ml of labeled *hRP S3a*-specific RNA at 68°C for 12 h. The blot was washed 3 × 10 min at 65°C with 0.5 × SSC/0.1% SDS. SSC is 0.15 M NaCl/0.015 M Na$_3$·citrate pH 7.6.
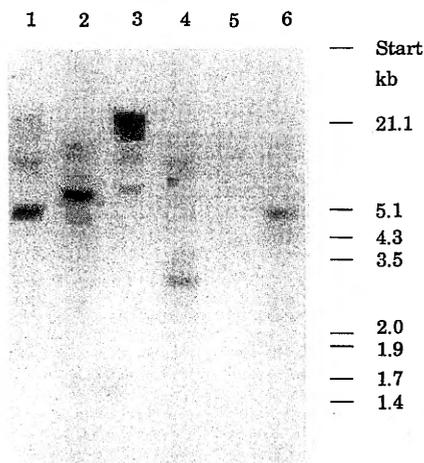
REFERENCES

Applied Biosystems User Bulletin No. 33 Foster City, CA, USA (1988).

Chen, I.-T. and Roufa, D. J.: The transcriptionally active human ribosomal protein S17 gene. Gene 70 (1988) 107–116.

Devi, K. R. G., Chan, Y.-L. and Wool, I. G.: The primary structure of rat ribosomal protein L18. DNA 7 (1988) 157–162.

Fischer, E. M. C., Beer-Romero, P., Brown, L. G., Ridley, A., McNeil, J. A., Lawrence, J. B., Willard, H. F., Bieber, F. R. and Page, D. C.: Homologous ribosomal protein genes of the human X and Y chromosomes: escape from X Inactivation and possible implications for Turner syndrome. Cell 3 (1990) 1205–1218.

Heinze, H., Arnold, H. H., Fischer, D. and Kruppa, J.: The primary structure of the human ribosomal protein S6 derived from a cloned cDNA. J. Biol. Chem. 263 (1988) 4139–4144.

Herzog, H., Höfferer, L., Schneider, R. and Schweiger, M.: cDNA encoding the human homologue of rat ribosomal protein L35a. Nucleic Acids Res. 18 (1990) 4600.

Laemmli, V. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (1970) 680–685.

316

Lastick, S. M. and McConkey, E. H.: Exchange and stability of HeLa ribosomal proteins in vivo. J. Biol. Chem. 251 (1976) 2867–2875.

Lott, J. B. and Mackie, G. A.: Isolation and characterization of cloned cDNAs that code for human ribosomal protein S6. Gene 65 (1988) 31–39.

Lutsch, G., Stahl, J., Kärgel, H.-J., Noll and F., Bielka, H.: Immunoelektron-microscopic studies on the location of ribosomal proteins on the surface of the 40S ribosomal subunit from rat liver. Eur. J. Cell. Biol. 51 (1990) 140–150.

Meyuhas O. and Perry, R. P.: Construction and identification of cDNA clones for mouse ribosomal proteins. Application for the study of r-protein gene expression. Gene 10 (1980) 113–129.

Nakamura, H., Tanaka, T. and Ishikawa, K.: Nucleotide sequence of cloned cDNA specific for rat ribosomal protein L7a. Nucleic Acids Res. 17 (1989) 4875.

Nobori, T., Hexdall, L. E. and Carson, D. A.: cDNA sequence of human ribosomal protein L31. Nucleic Acids Res. 17 (1989) 7105.

Rhoads, D. D., Dixit, A. and Roufa, D. J.: Primary structure of human ribosomal protein S14 and the gene that encodes it. Mol. Cell. Biol. 6 (1986) 2774–2783.

Rich, B. E. and Steitz, J. A.: Human acidic ribosomal phosphoproteins P0, P1 and P2: Analysis of cDNA clones, in vitro synthesis, and assembly. Mol. Cell. Biol. 7 (1987) 4065–4074.

Sambrook, J., Fritsch, E. F. and Maniatis, T.: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Shore D., Squire M. and Nasmyth, K. A.: Characterization of two genes required for the position-effect control of yeast mating-type genes. EMBO J. 3 (1984) 2817–2823.

Stanley, K. K. and Luzio, J. P.: Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. EMBO J. 3 (1984) 1429–1434.

Suzuki, K., Olvera, J. and Wool, I. G.: The primary structure of rat ribosomal protein L9. Gene 93 (1990) 297–300.

Tanaka, T., Kuwano, Y., Ishikawa, K. and Ogata, K.: Nucleotide sequence of cloned cDNA specific for rat ribosomal protein S11. J. Biol. Chem. 260 (1985) 6329–6333.

Tanaka, T., Wakasugi, K., Kuwano, Y., Ishikawa, K. and Ogata, K.: Nucleotide sequence of cloned cDNA specific for rat ribosomal protein L35a. Eur. J. Biochem. 154 (1986) 523–527.

Wettenhall, R. E. H., Nick, H. P. and Lithgow, T.: Primary structure of mammalian ribosomal protein S6. Biochemistry 27 (1988) 170–177.

Wool, I. G., Endo, Y., Chan, Y. and Glück, A.: Structure, function and evolution of mammalian ribosomes. In: Hill, W. E., Dahlberg, A., Garret, R. A., Moore, P. B., Schlessinger, D. and Warner, R. J. (Eds), The Ribosome. American Society for Microbiology, Washington, DC, 1990, pp. 203–214.

Zhang, X. T., Tan, Y.-M. and Tan, Y. H.: Isolation of a cDNA encoding human 40S ribosomal protein S3. Nucleic Acids Res. 18 (1990) 6689.

II

Rebane, A., Tamme, R., Laan, M., Pata, I. and Metspalu, A. (1998)
A novel snoRNA (U73) is encoded within the introns of the human and
mouse ribosomal protein S3a genes.
Gene 210, 255–263.

A novel snoRNA (U73) is encoded within the introns of the human and mouse ribosomal protein S3a genes

Ana Rebane [a], Richard Tamme [a,1], Maris Laan [b,2], Illar Pata [a], Andres Metspalu [a,c,*]

[a] Institute of Molecular and Cell Biology, Tartu University, Estonian Biocenter, 23 Riia St, EE2400 Tartu, Estonia
[b] Department of Clinical Chemistry, University of Helsinki, Helsinki, Finland
[c] Center of Molecular Diagnostics, Childrens Hospital, Tartu University, Oru 3, EE2400 Tartu, Estonia

Abstract

The mouse ribosomal protein S3a-encoding gene (mRPS3a) was cloned and sequenced in this study. mRPS3a shares identical exon/intron structure with its human counterpart. Both genes are split to six exons and exhibit remarkable conservation of the promoter region (68.8% identity in the 250 bp upstream of cap site) and coding region (the proteins differ in two amino acids). mRPS3a displays many features common to other r-protein genes, including the CpG-island at 5′-end of the gene, cap site within an oligopyrimidine tract and no consensus TATA or CAAT boxes. However, mRPS3a represents a rare subclass of r-protein genes that possess a long coding sequence in the first exon. Comparison of human and mouse S3a genes revealed sequence fragments with striking similarity within introns 3 and 4. Here we demonstrate that these sequences encode for a novel small nucleolar RNA (snoRNA) designated U73. U73 contains C, D and D′ boxes and a 12-nucleotide antisense complementarity to the 28S ribosomal RNA. These features place U73 into the family of intron-encoded antisense snoRNAs that guide site-specific 2′-O-ribose methylation of pre-rRNA. We propose that U73 is involved in methylation of the G1739 residue of the human 28S rRNA. In addition, we present the mapping of human ribosomal protein S3a gene (hRPS3a) and internally nested U73 gene to the human chromosome 4q31.2–3. © 1998 Elsevier Science B.V.

Keywords: Homo sapiens; Mus musculus; Polymorphism; Methylation; Chromosome 4q31.2–3

1. Introduction

Ribosomal protein S3a is located in the 40S ribosomal subunit in the region essential for initiation of translation. Immunoelectron microscopic and cross-linking

* Corresponding author. Tel: +372 7 420210; Fax +372 7 420286; e-mail: andres@ebc.ee
[1] Present address: Department of Genetics, University of Adelaide, Adelaide, SA 5005, Australia.
[2] Present address: Department of General Biology, Institute of Zoology, Luisenstr. 14, D-80333 München, Germany.

Abbreviations: aa, amino acid(s); FISH, fluorescence in situ hybridization; DAPI, 4′–6′ diamino-2-phenylindole; Gm, 2′-O-ribose methylated guanosine; ESTs, expressed sequence tags; L7a, gene encoding r-protein L7a; hRPS3a, gene encoding human S3a; mRPS3a, gene encoding mouse S3a; nt, nucleotide(s); PCR, polymerase chain reaction; pre-rRNA, preribosomal RNA; r-protein, ribosomal protein; S16, gene encoding r-protein S16; S3a, r-protein S3a; S3a, gene encoding S3a; snoRNA, small nucleolar RNA; SSC, 0.15 M NaCl/0.015 M Na$_3$·citrate (pH 7.6); tsp, transcription start point(s); U73, snoRNA U73; U73, gene encoding U73.

studies have shown that S3a can bind to eukaryotic initiation factors eIF-2 and eIF-3, as well as to initiator tRNA, to 18S rRNA and to mRNA (Westermann and Nygard, 1983; Lutsch et al., 1990; Bulygin et al., 1997). Consistent with its location, several studies indicate that expression of S3a gene and/or its homologues is required for cell growth and proliferation in various evolutionary distant species. Initially, the rat S3a gene was isolated as v-fos transformation effector gene (fte-1) (Kho and Zarbl, 1992). Later it was shown that monoallelic disruption of the fte-1 in v-fos transformed Rat-1 fibroblasts resulted not only in loss of the transformed phenotype, but also in decreased rate of protein synthesis (Kho et al., 1996). The gene for S3a homologue in Saccharomyces cerevisiae exists in duplicate copies (PLC1 and PLC2), and yeast spores carrying the disruption of either of these copies grow at slower rate than wild-type spores (Ito et al., 1992). In synchronously cultured Catharanthus roseus cells the cyc07 (encoding S3a homologue in plants) is preferentially and transiently expressed at the G$_1$/S boundary (Kodama et al.,

1991). The latter feature is characteristic of mammalian genes that are involved in regulation of the cell cycle and in DNA replication. The same behaviour has been suggested also for S3a and correspondingly presented as its possible extraribosomal function (Chan et al., 1996). In addition, rat r-protein S3b has the same aa sequence as rat S3a, except for missing the last 12 aa residues at carboxy-terminus. Although the origin of S3b remains unknown, researchers favour the idea that S3b results from proteolytic processing of S3a (Chan et al., 1996).

Cloning and sequencing of orthologous genes in two related species allows the use of homology analysis in order to identify functionally significant units within genes. It has now become evident that many snoRNAs are specified within introns of protein-encoding genes, rather than being transcribed independently (Smith and Steitz, 1997). A number of these genes encode either ribosome-associated or nucleolar proteins (reviewed by Maxwell and Fournier, 1995; Nicoloso et al., 1996). Since snoRNAs are located in the nucleolus, they are all expected to be involved in rRNA maturation in some manner. Therefore, it has been supposed that the preference of ribosome-related genes as parent genes for the snoRNAs is necessary in co-ordinating the production of protein and RNA components of the ribosome (Tycowski et al., 1993; Sollner-Webb, 1993).

Recently, substantial progress has been made in understanding the role of snoRNAs in rRNA maturation. Identification of novel conserved elements common to a group of snoRNAs (boxes ACA and H) has established a new class of snoRNAs, designated as the ACA family (Balakin et al., 1996; Ganot et al., 1997b). Latest reports suggest that ACA snoRNAs are involved in site-specific pseudo-uridine formation in pre-rRNA (Ni et al., 1997; Ganot et al., 1997a). Another and so far the largest group holds antisense snoRNAs characterized by sequence elements called box C (UGAUGA), box D/D′ (CUGA) and long stretches of complementarity to highly conserved rRNA sequences (Maxwell and Fournier, 1995). It has been elegantly demonstrated that antisense snoRNAs are required for site-specific 2′-*O*-ribose methylation of pre-rRNA and that the antisense elements, together with the D or D′ boxes, work as guides to select the target nucleotide for the methyl-transfer reaction (Kiss-László et al., 1996; Cavaillé et al., 1996).

In this study, we report the isolation and characterization of the gene encoding for the mouse S3a. We demonstrate that a novel snoRNA U73 is encoded within the introns of the human and mouse S3a gene. U73 possess C and D boxes, 12 nt long complementarity to the human and mouse 28S rRNA and most likely belongs to the class of snoRNAs that function in site-specific 2′-*O*-ribose methylation of RNA pre-RNA. We

also present the chromosomal sublocalization of human *S3a* and internally nested *U73* to 4q31.2–3.

2. Materials and methods

2.1. Isolation of lambda clones and DNA sequence analysis

A PCR-based screening method (Feo et al., 1990) was utilized to isolate the recombinant phage clones containing the mouse and human S3a encoding genes. The human lymphocyte genomic library in λDASH® vector (Stratagene, Cat. No. 943202) and the 129/Sv mouse genomic library in phage λFIXII vector (Stratagene, Cat. No. 946305) were divided into 36 pools, each containing 20 000 recombinant phage particles and amplified about 10^5-fold. The pools were screened with PCR using primer pairs derived from human S3a cDNA sequence (Metspalu et al., 1992). The primers 5′a (5′-CCTCTCCATGATGTCTTCG-3′) and 3′a (5′-GACCTTTTCGGTGACCCCTG-3′) were used for screening of the human genomic library, and 5′a and 3′b (5′-GACTACCTATACTTGGTGGTC-3′) for screening of the mouse genomic library. Primers were purchased from Bio-Synthesis (Lewisville, USA). PCR resulted in two amplification products: the larger corresponding to the intron-containing gene and the shorter to the processed pseudogenes. The pools positive for intron-containing gene were plated and screened using human and mouse ^{32}P-labelled S3a cDNA probes. Positive recombinant phages were purified to homogeneity. Various fragments (two *Eco*RI–*Bam*HI fragments of 3.0 kb and 3.3 kb for *hRPS3a* and two *Not*I fragments of 4.2 kb and 10.0 kb for *mRPS3a*) of λ clones were subcloned in the plasmid vector pBluescript II (Stratagene). DNA sequence analysis was performed by the dideoxy chain-termination method either using Sequenase 2.0 (USB) and [α-^{35}S]dATP or the Autoread 1000 sequencing kit (Pharmacia, Cat. No 27-1791-01) on ALF DNA Sequencer (Pharmacia). Sequence data were analysed with DNASIS and GeneSkipper computer programs. Other standard procedures were carried out as described earlier (Sambrook et al., 1989).

2.2. RNA analysis

The following oligodeoxynucleotides were used for Northern hybridization and primer extension analyses: for mouse and human U73a, m/hA3′ (5′-GTTATCA-GTGGCCATCAACTGGGA-3′); for mouse U73b, musB3′ (5′-AATGTCAGCAATAAATGTAACGGT-AA-3′); for human U73b, hB3′ (5′-ATAAATATAAT-GGTATAATCAGTATGT-3′); and for determining the *tsp* of *mRPS3a*, 5′-GTTCTTGCCGACCGCCATG-GTG-3′. Probes for Northern hybridization or primer

extension were 5′-end-labelled as follows. 10 pmol of oligodeoxynucleotide was incubated (45 min, 37°C) in the presence of 20 µCi [γ-³²P]ATP (5000 Ci/mmol, Amersham) and 10 units of T4 polynucleotide kinase (Fermentas, Lithuania), heat-inactivated (10 min, 75°C) and ethanol-precipitated.

Human cell lines PA-1, IGROV-1 (both from Glaxo Welcome, UK), Saos-2, HeLa and mouse C127 (ATCC) cells or intact mouse liver (female, strain C57BL/6, aged 6 weeks) were used for RNA analyses. Total RNA was extracted using RNAzol B solvent (WAK-Chemie, Germany) according to manufacturers' instructions.

For Northern analysis, 10 µg of total RNA was resolved on 8% acrylamide/7 M urea gel and electroblot-ted onto Hybond N⁺ membrane (Amersham, UK). After prehybridization in solution containing 6 × SSC, 1 × Denhardt, 0.5% SDS, 100 µg/ml denatured salmon sperm DNA (1 h at 65°C), hybridization was carried out in the fresh solution with 1 pmol of 5′-end-labelled oligodeoxynucleotide per ml for 5 h at 50°C. Filters were washed twice in 2 × SSC for 5 min at 50°C and exposed to X-ray film or analysed in PhosphoImager (Molecular Dynamics).

For primer extension analysis, 10 µg of total RNA was heat-denatured (10 min, 65°C) in the presence of 2–3 pmol end-labelled oligodeoxynucleotide and allowed to cool to room temperature. Reaction was carried out in the buffer containing 50 mM Tris–HCl (pH 8.3), 8 mM MgCl₂, 30 mM KCl, 1 mM each dNTP, 10 mM DTT and 20 units of AMV reverse transcriptase (Pharmacia) for 90 min at 42°C. Reaction was stopped by phenol–chloroform extraction, followed by alkaline denaturation and ethanol precipitation. The reaction products were resolved on a 6% sequencing gel and visualized by autoradiography. The secondary structure prediction for snoRNAs was performed with the RNADraw program.

2.3. Preparation of metaphase spreads and fluorescence in situ hybridization

To induce replication banding, phytohaemagglutinin-stimulated normal lymphocyte cultures were synchro-nized with 200 µg/ml 5-bromodeoxyuridine overnight, washed with PBS and reincubated with 0.3 µg/ml thymi-dime for additional 6–7 h (Lemieux et al., 1992; Tenhunen et al., 1995). Colcemid was added 20 min before harvesting. Metaphase spreads and microscope slides were prepared by standard protocols (Lemieux et al., 1992; Tenhunen et al., 1995). FISH was performed with a biotin–11-dUTP-labelled DNA probe derived from hRPS3a containing λ clone. Chromosomal banding was identified using the method for replication G-banding combined with 1% formaldehyde fixation of the metaphase preparations and staining with a low concentration of DAPI. The hybridization and visualiza-

tion procedures were carried out as described previously (Laan et al., 1995).

2.4. Database accession numbers

The complete nt sequences determined in this work have been deposited to EMBL, GenBank and DDBJ Nucleotide Sequence Databases under accession Nos (Z83368) mRPS3a, Z83334 (hRPS3a), Z83332 (mouse U73a), Z83331 (mouse U73b) and Z83330 (human U73a).

3. Results and discussion

3.1. Isolation and sequence analysis of mRPS3a and hRPS3a

Similarly to other mammalian r-protein genes, the human S3a forms a multigene family that consists of one functional intron-containing gene and about six to eight intronless, processed pseudogenes (Metspalu et al., 1992; Nolte et al., 1996). Southern analysis of mouse genomic DNA using S3a cDNA probe indicated that there are 9–11 members for S3a in the murine system (data not shown). To distinguish intron-containing genes from multiple pseudogenes during the isolation process of mRPS3a and hRPS3a, we used a combination of PCR and conventional library screening strategy as described by Feo et al. (1990). One λ clone containing genomic DNA for hRPS3a and one for mRPS3a were isolated, subcloned and the nt sequence was determined for both genes together with the flanking regions. The structural organization of hRPS3a and mRPS3a is pre-sented in Fig. 1, and the complete sequence for mRPS3a in Fig. 2. While the present work was in progress, Nolte
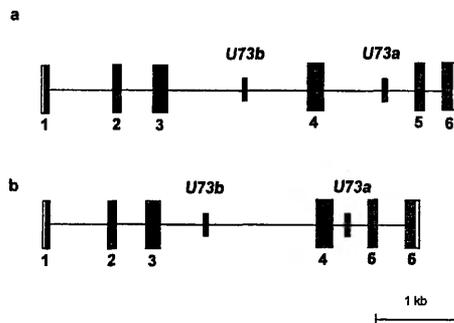

Fig. 1. Structure of hRPS3a (a) and mRPS3a (b) and locations of human and mouse U73 variants. The larger boxes designate exons (1–6), with black shading indicating coding sequences. The grey boxes show U73 positions within the hosts.

Fig. 2. Nucleotide sequence of the *mRPS3a* with predicted aa sequence. Exons are in upper-case and introns and flanking regions in lower-case letters. The transcription start point is marked by the caret (^); the stop codon is marked by the asterisk (*). The sequences matching with the putative transcription factor binding sites (Sp1, GABP, box A in human *L7a* and the relatively AT-rich area) and the polyadenylation sequence are underlined; sequences of *U73a* (intron 4) and *U73b* (intron 3) are underlined and in italics.

et al. (1996) reported the cloning and sequencing of human S3a-encoding gene. Comparison of the published *hRPS3a* sequence to that obtained in our laboratory (in the 5107 bp region) showed that both sequences are identical, except for 12 nt positions that all are located in introns. It remains unclear as to whether the differences between the two sequences represent the naturally occurring polymorphisms or sequencing errors. *mRRPS3a* is shorter than its human counterpart (4.6 and 5.1 kb, respectively). Both genes contain six exons (all of identical size) and five introns. Splice donor and acceptor sites follow the GT/AG rule (Mount, 1982). A computer search for repetitive sequences using the CENSOR server (censor@charon.girinst.org) identified a B2 repeat at nt positions 2358–2522; a RNALUIII between 1483 and 1549 and two B1 repetitive elements at nt positions 2768–2901 and 3187–3229 all located in the third intron of *mRPS3a*.

3.2. The mouse r-protein S3a

S3a homologues have been identified in 15 species, including vertebrates, plants, fungi and even archaebacteria (ProDom database at http://protein.toulouse. inra.fr/prodom.html). Although S3a protein has been highly conserved during evolution, it exhibits remarkable variability at the C-terminal region between aa 235 and 246 (Fig. 2). Similarly to mammalian homologues, the polypeptide deduced from the *mRPS3a* sequence consists of 264 aa. Compared to human S3a (Metspalu et al., 1992), it has two differences in this variable region: the Ser236 and Thr242 in human are replaced by Gly236 and Ala242 in mouse. We also performed computer analysis of mouse S3a cDNA sequences deposited in the database as ESTs (http://www.ncbi.nlm. nih.gov/dbEST/). Interestingly, we found that two aa from this region of S3a are polymorphic in laboratory mouse strains. Codon 237 can specify Ser (this study) or Thr (found in 68% of mouse S3a ESTs), and codon 238 may encode Ser (most frequent), as well as Arg (20%), Pro (5%) and Leu (2.5%).

3.3. Sequence analyses of the mRPS3a and hRPS3a promoter regions

Promoters of different r-protein genes share several features and transcription factor binding sites that can also be found in *mRPS3a* and *hRPS3a*. The main transcriptional start point *(tsp)* of *mRPS3a* was determined by primer extension analyses (data not shown) and it maps to the same nt (Fig. 2) within the oligopyrimidine tract as has been demonstrated for *hRPS3a* (Nolte et al., 1996). The first exon of *mRPS3a* and *hRPS3a* contains 25 nt of untranslated and 62 nt of coding sequence. This is atypical for vertebrate r-protein genes since ordinarily the first exon is either untranslated

or it encodes only the first few aa. Among all the mammalian r-protein genes characterized so far, only mouse *S16* has a similarly large first coding exon (Wagner and Perry, 1985). Typical to other housekeeping genes, the 5′ ends of *mRPS3a* and *hRPS3a* are embedded in a CpG-rich region (G+C content 64% and 60%, respectively, and CpG/GpC ratios 0.95). The promoters of both genes lack a well-defined TATA box, instead of which there is an A/T rich area between nt −28 and −22. Of putative binding sites for transcription factors, the *hRPS3a* and *mRPS3a* both contain two sites for GABP factor with consensus CGGAAR in either orientation (Hariharan et al., 1989) at nt positions −82 to −77 and −6 to −1, and two Sp1-recognition sites (Pugh and Tjian, 1990) at nt −3 to +3 and −42 to −37. In addition, *mRPS3a* has the sequence TCGCGAGA at nt −101 to −93, similar to the box A element in *L7a* promoter (Colombo and Fried, 1992) and three Sp1-binding sequences at nt positions +140 to +145, +164 to +169, +267 to 272 within the first intron that are not found in the human homologue (Fig. 2).

3.4. A novel snoRNA (U73) is specified within the introns of hRPS3a and mRPS3a

Sequence comparison of *hRPS3a* and *mRPS3a* revealed that introns 3 and 4 of both genes contain approx. 100 bp conserved elements with 72–85% identity to each other. Additionally, box C (UGAUGA) and box D (CUGA) motifs, hallmarks of several snoRNAs, were found in each element. The lack of significant sequence similarity of these conserved regions (except for C and D boxes) to any known snoRNAs prompted us to test by Northern blot hybridizations whether these snoRNA-like sequences encode novel snoRNA species. Using various antisense oligodeoxynucleotide probes, we were able to demonstrate that mouse introns 3 and 4 encode small stable RNAs of 73 and 69 nt, respectively. In human cell-lines HeLa, PA-1, IGROV-1 and Saos-2, we detected only the intron 4-encoded RNA of 65 nt (Fig. 3a). Henceforth we designated these RNAs as variants of a novel snoRNA U73: U73a (specified in intron 4) and U73b (in intron 3). The localizations of U73 genes within *hRPS3a* and *mRPS3a* introns are shown in Fig. 1 and alignment of human and mouse U73 sequences in Fig. 4a. The 5′ ends of the U73 snoRNAs were determined by primer extension analysis with various end-labelled oligodeoxynucleotides. As a result, the 5′ end for human U73a was mapped into nt position −6 relative box C, while for mouse U73a and U73b it was mapped into position −5 (Fig. 3aFig. 4a). Concordantly with the results of Northern analysis we did not detect specific primer extension product with oligos complementary to human U73b.
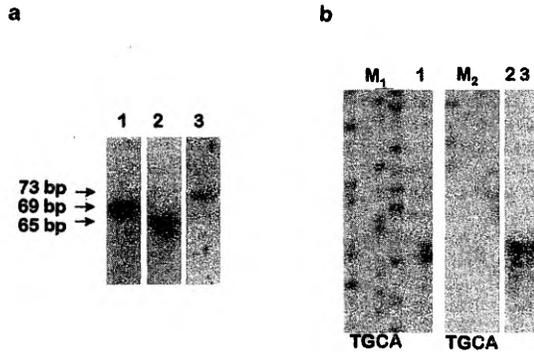
a   **b**

Fig. 3. (a) Detection of human and mouse U73 variants by Northern hybridization. Lanes 1, 2 and 3 correspond to mouse U73a (primer m/hA3′), human U73a (primer m/hA3′) and mouse U73b (primer musB3′). The sizes of snoRNAs were determined with respect to the RNase-T₁ ladder of *Escherichia coli* tRNA^{Phe} (Sigma) and M13 sequencing ladder and are indicated on the left. (b) The 5′ end mapping of human and mouse U73 expressed variants. The M_1 and M_2 are sequencing ladder size markers where the sequencing reaction was performed with primer specific to mouse U73b (musB3′). Lanes 1, 2 and 3 demonstrate the primer extension products for mouse U73b (primer musB3′), human U73a and mouse U73a (primer m/hA3′), the major extension products correspond to total lengths 70 bp for mouse U73b and 48 bp for mouse and human U73a (see also Fig. 4a). The oligonucleotides and conditions for both the Northern hybridization (a) and primer extension analyses (b) are described in Section 2.2.
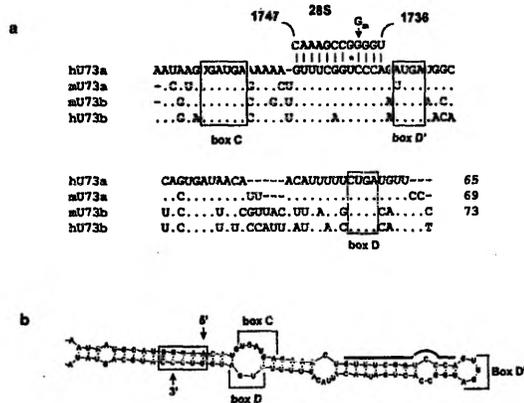

Fig. 4. (a) The sequence comparison of the human and the mouse U73 genes. Conserved boxes C, D′ and D are boxed and the complementary region of the human 28S rRNA (GenBank accession No. M11167) is shown above. Mouse 28S rRNA (GenBank accession No. X00525) has the respective sequence element between positions 1558 and 1570. The 5′ ends of mature snoRNAs were identified by primer extension and the 3′ ends were inferred from migration on denaturing polyacrylamide gels, making them accurate within one to two nucleotides. (b) The secondary structure model for mouse U73a together with flanking sequences from intronic RNA. The sequence tract complementary to 28S rRNA is overlined. The nt sequences of extended 5′–3′ terminal stem that are conserved between the expressed human and mouse U73 variants are boxed.

3.5. U73 belongs to the class of antisense snoRNAs and indicates a novel 2′-O-ribose methylation site in 28S rRNA

All the U73 variants in human and mouse possess boxes C, D and D′ (Fig. 4a) characteristic of the antisense snoRNAs that participate in 2′-O-ribose methyla-tion of rRNAs, leading to search for the potential base-pairing between U73 and human/mouse pre-rRNA. We found that a 12-nt sequence in U73 just upstream of the D′ box is complementary to a highly conserved region both in the human and mouse 28S rRNAs (including one G–U pairing). The sequence of this antisense element is 100% identical in all the expressed variants of
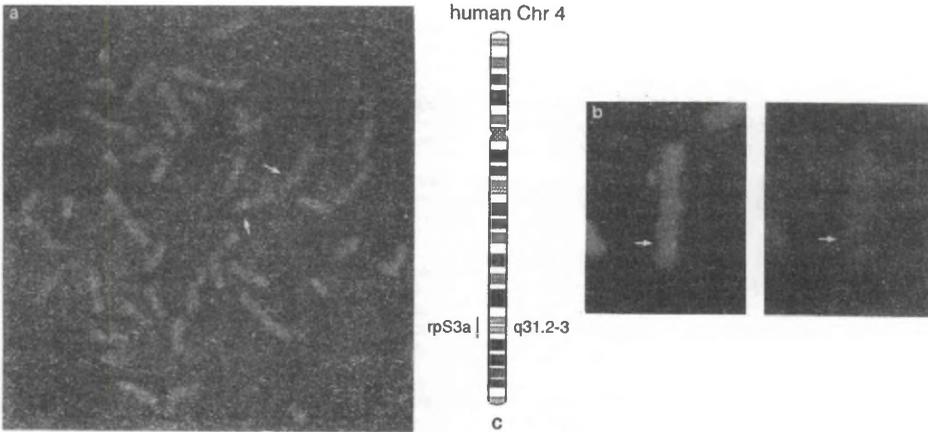
human Chr 4

rpS3a | q31.2-3

c

Fig. 5. Localization of the *hRPS3a* and internally encoded *U73* gene using FISH. (a) Partial human metaphase spread counterstained with propidium iodide, showing specific fluorescent hybridization signals on two chromosomes 4. (b) Enlarged chromosome 4 revealing the localization of *hRPS3a* to band 4q31.2–3. Assignment of the locus was performed using parallel examination of intensified DAPI-banding. (c) Schematic representation of human chromosome 4 with the hybridization locus of *hRPS3a* indicated.

U73. By contrast, the non-expressed human variant U73b has a G–A nucleotide substitution in this region that alters the snoRNA–rRNA pairing potential (Fig. 3a).

It has been demonstrated that the fifth nt upstream of the D or D' box in the antisense element of the snoRNAs always faces the methylated nucleotide in the rRNA sequence (Kiss-László et al., 1996; Cavaillé et al., 1996). The fifth nt upstream of the D' box of U73 corresponds to G1739 in human 28S rRNA (nt numbering according to GenBank accession No. M11167). Similarly to ribose-methylated residues, the G1739 resides in a bulge region of a phylogenetically conserved loop of 28S rRNA. None of the 2'-O-methyl groups determined so far has been placed in this position or in the close vicinity (Maden, 1990). However, about 11–13 2'-O-ribose methylated nucleotides in mammalian 28S rRNA remain to be mapped precisely. Among these unplaced nucleotides, at least seven methylated G residues occur in the sequence context G–Gm–G (Maden, 1990). Concordantly, the G1739 is located in the midst of the sequence element GGG. Therefore, these data together with the similarity of U73 to other snoRNAs that have been implicated in rRNA ribose methylation, strongly suggest that U73 may guide 2'-O-ribose methylation of G1739 in the human 28S RNA.

3.6. The secondary structure of U73 snoRNA

The secondary structures for the human and mouse U73 variants were predicted on a thermodynamic basis using the RNADraw program. The mouse and human U73 variants (except for the non-expressed human U73b) form two stem structures as described for several other intronic C/D box containing snoRNAs (Tycowski et al., 1993; Nicoloso et al., 1994, 1996). A long but imperfect stem between the C and D boxes folds the U73 variants into a hairpin structure, so that the C and D box motifs are juxtaposed. Another 3–4 bp-long stem holds together the 5' and 3' ends of U73 (Fig. 4b). The intronic sequences flanking the coding regions of expressed U73 variants are also able to pair and to form the stem extension. Interestingly, the crossing area of terminal stem/flanking stem is conserved so that the 5' and 3' ends of human and mouse variants map into the same sequence element (GGGAA and UUCCC, respectively; Fig. 4b). By contrast, our experiments have shown that the 5' and 3' end positions of different U73 variants within these elements are not conserved. It seems that primary sequence conservation in this region is required for forming the terminal (and flanking) stem, rather than determining the 5' and 3' ends. This finding supports the current model according to which the juxtaposition of C and D boxes by terminal and/or internal stem controls the binding of protein(s) that protect mature snoRNA against exonucleases (Xia et al., 1997).

3.7. Chromosomal localization of hRPS3a and the internally nested U73

Using fluorescent microscopy, 35 metaphase spreads hybridized with the *hRPS3a* genomic probe were exam-

18

ined. Fluorescent hybridization signals were seen on one or both homologues of chromosome 4 in 89% of metaphases (Fig. 5a). No other locus with specific fluorescent spots was detected. Parallel examination of the hybridization signals and the DAPI-banded chromosomes assigned the locus for *hRPS3a* and internally nested *U73* to the long arm of chromosome 4 to the band q31.2–3 (Fig. 5b,c).

4. Conclusions

(1) *mRPS3a* and *hRPS3a* share identical exon/intron structure, exhibit remarkable conservation in the promoter region, and represent a rare subclass of r-protein genes that possess a long coding sequence in the first exon.

(2) The polypeptide deduced from *mRPS3a* has two aa replacements relative to the human S3a protein, both of which are located in a variable region close to the S3a C-terminus. Two other aa in this region are polymorphic in laboratory mice.

(3) Introns 3 and 4 of *hRPS3a* and *mRPS3a* encode a novel snoRNA termed U73. Both variants of *U73* are expressed in mouse, while only one variant (specified by intron 4) is expressed in tested human cell-lines.

(4) U73 contains features similar to intron-encoded antisense snoRNAs that guide site-specific $2'$-O-ribose methylation of pre-rRNA. We propose that the U73 is involved in methylation of G1739 residue of the human 28S rRNA.

(5) *hRPS3a* and internally encoded U73 genes map to the human chromosome 4q31.2–3.

Acknowledgement

References

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

Bulygin, K.N., Matasova, N.B., Graifer, D.M., Veniyaminova, A.G., Yamkovoy, V.I., Stahl, J., Karpova, G.G., 1997. Protein enviroment of mRNA at the decoding site of 80S ribosomes from human placenta as revealed from affinity labeling with mRNA analogs – derivatives of oligoribonucleotides. Biochim. Biophys. Acta 1351, 325–332.

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.

Chan, Y.L., Olvera, J., Paz, V., Wool, I.G., 1996. The primary structures of rat ribosomal proteins S3a (the V-Fos transformation effector) and of S3b. Biochem. Biophys. Res. Commun. 228, 141–147.

Colombo, P., Fried, M., 1992. Functional elements of the ribosomal protein L7a (rpL7a) gene promoter region and their conservation between mammals and birds. Nucleic Acids Res. 20, 3367–3373.

Feo, S., Davies, B., Fried, M., 1990. A phage screening method to isolate intron-containing genes in presence of multiple processed pseudogenes. Nucleic Acids Res. 16, 4949–4950.

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 evolutionary conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. Genes Dev. 11, 941–956.

Hariharan, N., Kelley, D.E., Perry, R.P., 1989. Equipotent mouse ribosomal protein promoters have similar architecture that includes internal sequence elements. Genes Dev. 3, 1789–1800.

Ito, M., Yasui, A., Komamine, A., 1992. A gene family homologues to the S-phase specific gene in higher plants is essential for cell proliferation in *Saccharomyces cerevisiae*. FEBS Lett. 301, 29–33.

Kho, C.J., Wang, Y., Zarbl, H., 1996. Effect of decreased *fte-1* gene expression on protein synthesis, cell growth, and transformation. Cell Growth Differ. 7, 1157–1166.

Kho, J.K., Zarbl, H., 1992. *Fte-1*, a v-*fos* transformation effector gene, encodes the mammalian homologue of a yeast gene involved in protein import into mitochondria. Proc. Natl. Acad. Sci. USA 89, 2200–2204.

Kiss-László, Z., Yves, H., 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.

Kodama, I., Ito, M., Hattori, T., Nakamura, K., Komamine, A., 1991. Isolation of genes that are preferentially expressed at the G_1/S boundary during the cell cycle in synchronized cultures of *Chataranthus roseus* cells. Plant Physiol. 95, 406–411.

Laan, M., Grön-Virta, K., Salo, A., Aula, P., Peltonen, L., Palotie, A., Syvänen, A.C., 1995. Solid-phase minisequencing confirmed by FISH analysis in determination of gene copy number. Hum. Genet. 96, 275–280.

Lemieux, N., Dutrillaux, B., Viegas-Péquignot, E., 1992. A simple method for simultaneous R- or G-banding and fluorescence in situ hybridization of small single-copy genes. Cytogenet. Cell Genet. 59, 311–312.

Lutsch, G., Stahl, J., Kärgel, H.J., Bielka, H., 1990. Immunoelectron microscopic studies on the location of ribosomal proteins on the surface of 40S ribosomal subunit from rat liver. Eur. J. Cell Biol. 51, 140–150.

Maden, B.E.H., 1990. The numerous modified nucleotides in eukaryotic ribosomal RNA. Prog. Nucl. Acids Res. Mol. Biol. 39, 241–303.

Maxwell, E.S., Fournier, M.J., 1995. The small nucleolar RNAs. Annu. Rev. Biochem. 35, 897–934.

Metspalu, A., Rebane, A., Hoth, S., Pooga, M., Stahl, J., Kruppa, J., 1992. Human ribosomal protein S3a: cloning of the cDNA and primary structure of the protein. Gene 119, 313–316.

Mount, S.M., 1982. A catalogue of splice junction sequences. Nucleic Acids Res. 10, 459–472.

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., Caizergues-Ferrer, M., Michot, B., Azum, M.-C., Bachellerie, J.-P., 1994. U20, a novel small nucleolar RNA, is encoded in an intron of the nucleolin gene in mammals. Mol. Cell. Biol. 14, 5766–5776.

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

Nolte, D., Taimor, G., Kalf-Suske, M., Seifart, K.H., 1996. The human S3a ribosomal protein: sequence, localisation and cell-free transcription of the functional gene. Gene 169, 179–185.

Pugh, B.F., Tjian, R., 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell 61, 118711–118797.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Laboratory Press, Cold Spring Harbor, NY.

Smith, C.M., Steitz, J.A., 1997. Sno storm in the nucleolus: new roles for muriad small RNPs. Cell 89, 669–672.

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

Tenhunen, K., Laan, M., Manninen, T., Peltonen, L., Jalanko, A., 1995. Molecular cloning, chromosomal assignment, and expression of the mouse aspartylglucosaminidase gene. Genomics 30, 244–250.

Tycowski, K.T., Shu, M.-D., Steitz, J.A., 1993. A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. Gen. Dev. 7, 1176–1190.

Wagner, M., Perry, R.P., Characterization of the multigene family encoding the mouse S16 ribosomal protein: strategy for distinguishing an expressed gene from its processed pseudogene counterparts by an analysis of total genomic DNA. 1985. Mol. Cell. Biol. 5, 3560–3576.

Westermann, P., Nygard, O., 1983. The spatial arrangement of the complex between eukaryotic initiation factor eIF-3 and 40 S ribosomal subunit. Cross-linking between factor and ribosomal protein. Bichim. Biophys. Acta 741, 103–108.

Xia, L., Watkins, N.J., Maxwell, E.S., 1997. Identification of specific nucleotide sequences and structural elements required for intronic U14 snoRNA processing. RNA 3, 17–26.

III

Rebane, A. and Metspalu A. (1999)
U82, a novel snoRNA identified from the fifth intron of human and
mouse nucleolin gene.
Biochim. Biophys. Acta 1446, 426–430.

Short sequence-paper

U82, a novel snoRNA identified from the fifth intron of human and mouse nucleolin gene[1]

Ana Rebane [a], Andres Metspalu [a,b,*]

[a] *Institute of Molecular and Cell Biology, Tartu University, Estonian Biocentre, 23 Riia St., 51010 Tartu, Estonia*
[b] *Centre of Molecular Diagnostics, Children's Hospital, Tartu University, Oru 3, Tartu, 51005, Estonia*

Abstract

A novel snoRNA, designated as U82, was found from the sequence analysis of the 5th intron of human and mouse nucleolin gene. The snoRNA U82 has characteristic boxes C, D and D' and 11 nucleotides (nt) antisense complementarity to the 18S rRNA. Presumably U82 functions as a guide in the methylation of residue A1678 in human 18S rRNA. Northern blot analysis with various oligodeoxynucleotide probes showed that human and mouse U82 is expressed as RNA variants with length of 70 (± 1) and 67 (± 1) nt in HeLa and mouse C127 cells. Most probably, the 70 nt variant of U82 is encoded by nucleolin gene 5th intron. The 67 nt variant of U82 could be a transcript of another gene, the genomic locus of which remains unknown.   © 1999 Elsevier Science B.V. All rights reserved.

Keywords: snoRNA; U82; Nucleolin; 2'-*O*-Methylation

Nucleoli of eukaryotic cells contain a large number of distinct small RNAs, termed as small nucleolar RNAs (snoRNAs) [1,2]. The majority of snoRNAs function as guide RNAs in the nucleotide modification of preribosomal RNA (pre-rRNA). A large group of these snoRNAs is defined by sequence elements as box C (UGAUGA) and box D (CUGA) [3,4]. It has been demonstrated that C/D box snoRNAs take part in site-specific 2'-*O*-ribose methylation of pre-rRNA [5,6]. Another class of modification snoRNAs share conserved sequence elements H (ANANNA) and ACA [7,8]. The H/ACA sno-RNAs have been shown to direct the site-specific synthesis of pseudouridine in pre-rRNA [9,10]. In vertebrates, all the known pre-rRNA modification snoRNAs are encoded within introns of protein-coding genes or processed from pre-mRNA-like primary transcripts which do not generate functional mRNAs [2,11–13]. All established snoRNA host genes belong to the 5'-terminal oligopyrimidine gene family [13,14]. A number of these genes encode proteins that are involved in nucleolar function, ribosome structure or protein synthesis [1–3,14]. In this study, we report the identification of a novel C/D box snoRNA from the fifth intron of human and mouse nucleolin gene.

In order to identify new snoRNA-encoding sequences, the introns of orthologous ribosome-associated genes were compared to each other using Fasta and PlotSimilarity programs of the GCG package.

* Corresponding author. Fax: +372-7-420286;
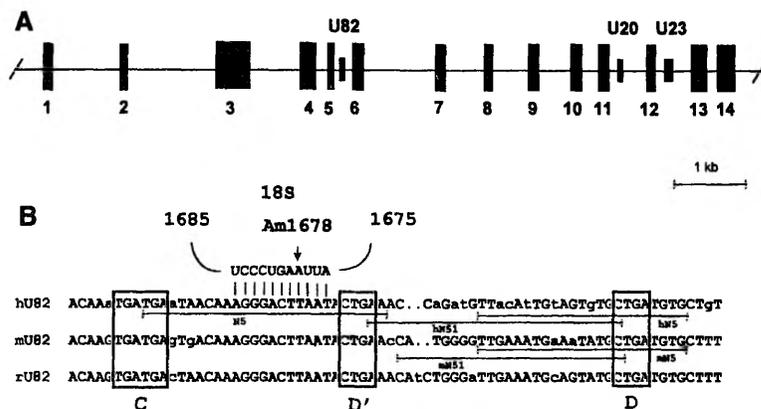E-mail: andres@ebc.ee

Fig. 1. (A) The structure of human nucleolin gene and locations of internally nested snoRNA genes. The boxes with grey shading designate exons. The black boxes show positions of snoRNA genes within the host gene. (B) The sequence comparison of human, mouse and rat U82 genes. Conserved boxes C, D′ and D are indicated and the complementary region of human 18S rRNA (GenBank accession no. X03205) is shown above. Location of oligodeoxynucleotide probes used in Northern hybridization is shown by bars below the sequence. The nucleotides that do not agree with consensus are shown in lowercase. The expression, length and 5′ end position of rat U82 was not determined experimentally.

Interestingly, a highly conserved snoRNA-like region that has no significantly similar sequence (except C, D and D′ boxes) to any known snoRNA was found in the 5th intron of human, rat and mouse nucleolin gene. Earlier, was shown that mammalian nucleolin gene intron 11 encodes C/D box snoRNA U20 [15]. Intron 12 of human, mouse and rat nucleolin gene has been known to specify H/ACA snoRNA U23 gene (GenBank accession no. AJ007015). The structure of human nucleolin gene and positions of internally nested snoRNA genes are presented in Fig. 1A. The sequence comparison of human, rat and mouse U82 is shown in Fig. 1B. As snoRNAs up to U81 have been reported previously [13], the conserved snoRNA-like region of nucleolin gene 5th intron is designated as U82 coding gene.

To identify the expression of U82 snoRNA, Northern blot analysis of RNA extracted from human HeLa and mouse C127 cells (ATCC) was carried out. Following oligodeoxynucleotides (Genset) were used as probes. hN5 (5′ CAC ATC AGC ACA CTA CAA TGT AA 3′) and hN51 (5′ GCA CAC TAC AAT GTA ACA TCT GGT TT 3′) were designed according to human nucleolin gene sequence (GenBank accession no. M60858), mN5 (5′

CAC ATC AGC ATA TTT CAT TTC AA 3′) and mN51 (5′ GCA TAT TTC ATT TCA ACC CCA TG 3′) according to mouse nucleolin gene sequence (GenBank accession no. X07699). N5 (5′ TTC AGT ATT AAG TCC CTT TGT TAC TCA 3′) was designed to agree with one mismatch with the both. Procedures for 5′ end-labelling of the oligodeoxynucleotide probes and RNA extraction were performed as previously reported [16]. For Northern analysis, 10 μg of total cellular RNA was resolved on 10% acrylamide/7 M urea gel and electroblotted onto Hybond N+ membrane (Amersham). 100 pg of in vitro transcript of cloned human nucleolin gene 5th intron as a stringency reference was added in each lane. To generate the in vitro transcript, a 180 bp fragment of human nucleolin gene 5th intron was amplified with PCR and cloned into the BamHI–PstI sites of pBluescript SK⁻ (Stratagene), resulting in clone N5PCR. In vitro RNA synthesis was carried out on EcoRI-linearized N5PCR, using T3 RNA polymerase (Fermentas) according to manufacturer's instructions. Prehybridization and Northern hybridization were carried out as reported [16], except that the hybridization was performed at 50°C for 3 h. Filters were washed at least twice in 2×SSC for 10 min at 45°C;
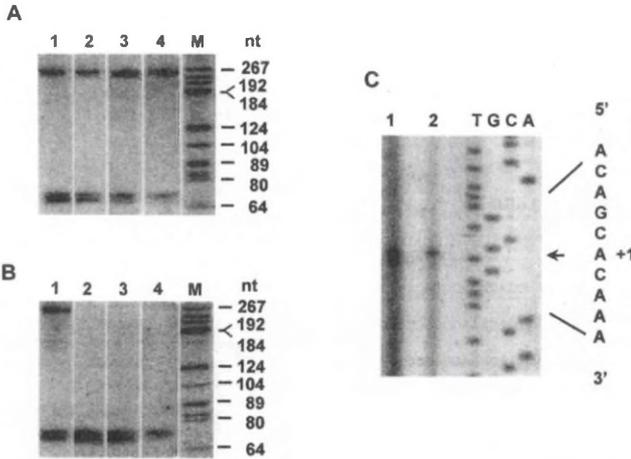
Fig. 2. Expression and mapping the 5′ ends of human and mouse U82 snoRNAs. (A) Northern blot analysis of human U82. Each lane (1–4) contains 10 μg of total RNA extracted from HeLa cells and 100 pg in vitro transcript of 240 nt of cloned human nucleolin gene 5th intron. Oligodeoxynucleotides N5 (lane 1), hN51 (lane 2) and hN5 (lanes 3 and 4) were used as probes. Filters were washed at 45°C (lanes 1–3) or 55°C (lane 4). (B) Northern hybridization of mouse U82. Lanes 1–4 contain 10 μg of total RNA extracted mouse C127 cells and 100 pg in vitro transcript of 240 nt of cloned human nucleolin gene 5th intron. Oligodeoxynucleotide probes N5 (lane 1), mN51 (lane 2) and mN5 (lanes 3 and 4) were used. Filters were washed at 45°C (lanes 1–3) or at 52°C (lane 4). Lanes designated with M present BsuRI-digested and 5′ end-labelled pBR322 DNA as marker. (C) Mapping the 5′ ends of human and mouse U82. Lanes 1 and 2 demonstrate the primer extension analysis of total RNA from mouse C127 and human HeLa cells with probes mN5 and probe hN5, respectively. Lanes TCGA show sequencing ladder generated on clone N5PCR with primer hN5. The position of main primer extension products is indicated by an arrow.

subsequently some of filters were washed in 2×SSC at various higher temperatures from 48°C to 60°C. As a result, when filters were washed at 45°C, two RNAs approximately 70 nt and 67 nt were detected from HeLa and mouse C127 cells with all oligodeoxynucleotide probes (Fig. 2A,B). The length of these two RNAs was determined in denaturing polyacrylamide gels and it is accurate within one or two nucleotides. Subsequently, the snoRNA lengths of 70 nt and 67 nt were used in this study. As expected, the mouse-specific probes mN5 and mN51 did not hybridize to in vitro synthesized RNA transcript the sequence of which corresponds to human nucleolin gene 5th intron. Interestingly, with the oligodeoxynucleotide probes hN5 and mN5, only the 70 nt band of RNA was visible with washing temperatures 55°C and 52°C, respectively (Fig. 2A,B). With probe hN5, the in vitro synthesized RNA band was always clearly seen together with the 70 nt RNA band even

in conditions where the 67 nt band disappeared With the oligodeoxynucleotide probes N5, hN51 and mN51 both RNAs could be detected either with the same intensity or not at all. Thus, Northern blot analysis showed that snoRNA U82 is expressed as two length variants of 70 nt and 67 nt in HeLa and mouse C127 cells. Absence of the 67 nt RNA band at stringent washing conditions with probes hN5 and mN5 points to internal sequence differences between 70 nt and 67 nt species. Therefore, it is likely that U82 variants of 70 nt and 67 nt are not two different processing products of the same transcript, but rather the products of different genes. Most probably, the 70 nt variant of U82 is encoded by nucleolin gene 5th intron. The 67 nt variant of U82 could be specified by another gene, the genomic locus of which remains unknown.

The 5′ ends of the human and mouse U82 snoRNAs were determined by primer extension analysis

using oligodeoxynucleotides hN5 and mN5, respectively. The mapping procedure was carried out as described earlier [16]. The sequencing ladder as a marker was generated with the primer hN5 on clone N5PCR, the latter being produced as described above. As a result, the 5′ ends of human and mouse U82 were mapped to the nt position −5 from the box C (Fig. 1B and Fig. 2C).

All the C/D box snoRNAs described so far contain antisense element(s) of 10–21 nt complementary to pre-rRNA [17]. Using the Findpattern program of the GCG package, potential base-pairing between U82 and human pre-rRNA was investigated. It was found that the 11 nt region just upstream the D′ box of U82 was complementary to the sequence between 1675 and 1686 nt of human 18S rRNA (nt numbering according to GenBank accession no. X03205) (Fig. 1B). The fifth nucleotide upstream the D or D′ box in the antisense element of the snoRNA always faces the methylated nucleotide in the rRNA sequence [4–6,17]. In U82, the fifth nucleotide upstream the D′ box corresponds to A1678 in human 18S rRNA (Fig. 1B). Accordingly, A1678 is determined to be one of 40 2′-*O*-methylated nucleotides in human 18S rRNA [18]. The 2′-*O*-methylated residue A1678 in human 18S rRNA is conserved in *Xenopus laevis* as well, but not in yeast [18].

In conclusion, a novel snoRNA U82 was identified from human and mouse nucleolin gene in addition to two previously known snoRNAs, U20 and U23. U82 contains features similar to snoRNAs that guide site-specific 2′-*O*-ribose methylation of pre-rRNA and presumably takes part in methyl-transfer reaction of A1678 in human 18S rRNA. Northern blot analysis has shown that U82 is expressed as two variants with estimated lengths of 70 nt and 67 nt in HeLa and mouse C127 cells. The 67 nt variant of U82 could be specified by another locus than the 5th intron of nucleolin gene.

After this manuscript was first submitted, the sequences matching with human, mouse and rat U82 gene were released in EMBL/GenBank/DDBJ databases as human, mouse and rat Z25 genes under the accession numbers AJ01066, AJ01067 and AJ010668, respectively.

References

[1] C.M. Smith, J.A. Steitz, Sno storm in the nucleolus: new roles for muriad small RNPs, Cell 89 (1997) 669–672.

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

[3] E.S. Maxwell, M.J. Fournier, The small nucleolar RNAs, Annu. Rev. Biochem. 35 (1995) 897–934.

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

[5] Z. Kiss-László, H. Yves, J.-P. Bachellerie, M. Caizergues-Ferrer, T. Kiss, Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs, Cell 85 (1996) 1077–1088.

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

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

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

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

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

[11] K.T. Tycowski, M.-D. Shu, J.A. Steitz, A mammalian gene with introns instead of exons generating stable RNA products, Nature 379 (1996) 3158–3166.

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

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

[14] P. Pelczar, W. Filipowicz, The host gene for intronic U17

small nucleolar RNAs in mammal has no protein-coding potential and is a member of the 5′-terminal oligopyrimidine gene family, Mol. Cell. Biol. 18 (1998) 4509–4518.

[15] M. Nicoloso, M. Caizergues-Ferrer, B. Michot, M.-C. Azum, J.-P. Bachellerie, U20, a novel small nucleolar RNA, is encoded in an intron of the nucleolin gene in mammals, Mol. Cell. Biol. 14 (1994) 5766–5776.

[16] A. Rebane, R. Tamme, M. Laan, I. Pata, A. Metspalu, A novel snoRNA (U73) is encoded within the introns of the human and mouse ribosomal protein S3a genes, Gene 210 (1998) 255–263.

[17] J.-P. Bachellerie, J. Cavaillé, Small nucleolar RNAs guide the ribose methylations of eukaryotic rRNAs, in: H. Grosjean, R. Benne (Eds.), Modification and Editing of RNA, ASM Press, Washington, 1998, pp. 255–272.

[18] B.E.H. Maden, The numerous modified nucleotides in eukaryotic ribosomal RNA, Prog. Nucleic Acids Res. Mol. Biol. 39 (1990) 241–303.

CURRICULUM VITAE

ANA REBANE

Date of birth: August 11, 1966
Citizenship: Estonian
Family status: married, two children
Work address: Institute of Molecular and Cell Biology,
                Tartu University,
                Riia Str. 23, Tartu 51010, Estonia
Phone: 372 7 375034; 372 7 375029
Fax: 372 7 420286
E-mail: arebane@ebc.ee
Home address: Elva Str. 43–4, Tartu 50411, Estonia

Education and professional employment

1984   — finished Tartu Secondary School No 7
1989   — graduated from Tartu University, Estonia; Faculty of Physics
            and Chemistry, as bioorganic chemist
1989–1995 — Estonian Biocentre, junior research scientist
1995–1999 — Tartu University, Faculty of Biology and Geography, Institute
            of Molecular and Cell Biology, Ph. D. student
1996–1997 — working and studying in A. I. Virtanen Institute, University of
            Kuopio (Finland) with scholarships from Nordic Council of
            Ministers and Finnish Government during 11 months

Scientific work

In 1989 I started to work on mammalian ribosomal protein genes under the
supervision of Dr. Andres Metspalu. Currently I am studying small nucleolar
RNAs.

CURRICULUM VITAE

ANA REBANE

Sünniaeg:	11. august 1966
Kodakondsus:	Eesti
Perekonnaseis:	abielus, kaks last
Aadress tööl:	Tartu Ülikooli molekulaar- ja rakubioloogia instituut Riia t. 23, Tartu 51010
Tel.:	27 375034; 27 375029
Faks:	27 420286
E-mail:	arebane@ebc.ee
Kodune aadress:	Elva t. 43–4, Tartu 50411

Haridus ja erialane teenistuskäik

1981–1984 — Tartu 7. Keskkool

1984–1989 — Tartu Ülikooli füüsika-keemiateaduskond, bioorgaanilise keemia eriala üliõpilane

1989–1995 — Eesti TA, Eesti Biokeskus, nooremteadur

1995–1999 — Tartu Ülikool, bioloogia-geograafiateaduskond, molekulaar- ja rakubioloogia instituut, doktorant

1996–1997 — õppimine ja töötamine Kuopio Ülikooli A. I. Virtaneni instituudis, kokku 11 kuud, Põhjamaade Ministrite Nõukogu ja Soome valitsuse stipendiumidega

Teadustegevus

Alates 1989. aastast olen uurinud dr. Andres Metspalu juhendamisel imetajate ribosoomivalkude geene ning viimasel ajal ka väikesi tuumakese RNAsid

DISSERTATIONES BIOLOGICAE
UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida*. Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Liliam Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand**. The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.

19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.

21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.

22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.

23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.

24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.

25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.

26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.

27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.

28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.

29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.

30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.

31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.

32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.

33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.

34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.

35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.

36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.

37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.

38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.

39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.

40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.
41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.